Univerzita Karlova v Praze

Přírodovědecká fakulta

Biologie

Buněčná a vývojová biologie

The role of acetylation in the RNA recognition motif of SRSF5 protein

Role acetylace RNA vazebného motivu proteinu SRSF5

Diplomová práce

Bc. Jaroslav Icha

školitel: Mgr. David Staněk, Ph.D.

Praha, 2012

Prohlašuji, že jsem závěrečnou práci zpracoval samostatně a že jsem uvedl všechny použité informační zdroje a literaturu. Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu.

V Praze, 24.8.2012

Acknowledgments

I thank to my supervisor, David Staněk, for guidance during my MSc. studies. I am thankful to Michaela Blazikova for help with FRAP data analysis, Václav Urban and Igor Shevelev for assistance during protein purification and Jana Písačková for help with thermofluor measurement. I want to emphasize the role of Veronika Bačíková from Richard Štefl`s lab, who performed the fluorescence anizotropy measurement. I am grateful to Eva Dušková for sharing the fibronectin minigene reporter, Karla Neugebauer for gifts of cell lines and antibodies and all the Staněk lab members for valuable discussions and pleasant working environment. This work was supported by the Czech Science Foundation grant no.: P305/10/0424; 2010-2013: Regulation of alternative splicing via chromatin acetylation.

Objective of the thesis

The primary objective of this thesis is to explore the function of acetylation of SRSF5 protein by characterizing the phenotype of SRSF5 proteins with mutated acetylation site. The hypothesis to be tested is that the properties of amino acid side chain at the site of acetylation influence interaction of SRSF5 with RNA.

Abstract

Acetylation is emerging as an important posttranslational modification, which is found in thousands of proteins in eukaryotes, as well as prokaryotes. Global proteomic studies implicated acetylation in regulation of various processes like metabolism, gene expression, cell cycle or aging to name a few. In this work I set out to investigate the role of acetylation of a splicing regulatory protein SRSF5 by creating mutations in its acetylation site. I tested the hypothesis that acetylation influences SRSF5 interaction with RNA. I expressed acetylation-mimicking (Q) or non-acetylable (R) mutant of SRSF5 in HeLa cells and measured their interaction with RNA by RNA immunoprecipitation or *in vitro* by fluorescence anisotropy. Both approaches agreed that mutants interact with RNA less than the wild type protein and Q mutant bound RNA weaker than R mutant.

I did not detect further difference in localization or dynamics among the proteins *in vivo*, which suggests that difference caused by weakened interaction of mutants with RNA was outweighed by other factors influencing SRSF5 behaviour, probably protein-protein interactions. I also found out that mutant SRSF5 proteins do not have a dominant effect on splicing of fibronectin alternative EDB exon.

The data obtained give an indirect evidence for the hypothesis that acetylation influences binding of SRSF5 to RNA. The low RNA affinity of acetylation-mimicking mutant suggests that acetylation reduces SRSF5 binding to RNA. This finding might be intersting, because there is a lot of other proteins with reported acetylation in RNA-binding domain implicating acetylation in regulation of RNA binding and consequently in regulation of pre-mRNA splicing.

key words: acetylation, SRSF5, RNA binding, RNA-protein interaction, SR proteins, RNA splicing

Abstrakt

Acetylace je významná posttranslační modifikace, která byla nalezena u několika tisíc proteinů jak u eukaryot, tak u prokaryot. Globální proteomické studie odhalily, že acetylace reguluje mnoho procesů např. metabolizmus, genovou expresi, buněčný cyklus, stárnutí a mnohé další. V této práci jsem studoval roli acetylace sestřihového faktoru SRSF5 pomocí mutací v jeho acetylačním místě. Testoval jsem hypotézu, že acetylace ovlivňuje vazbu SRSF5 na RNA. V HeLa buňkách jsem exprimoval mutované proteiny SRSF5, jednak mutanta napodobujícího acetylovanou formu (Q) a také neacetylovatelného mutanta (R) a zjišťoval jsem jejich interakci s RNA pomocí RNA imunoprecipitace a také *in vitro* měřením fluorescenční anizotropie. Výsledky obou metod se shodly na tom, že oba mutanty interagují s RNA méně než přirozená forma proteinu a Q mutant se vázal na RNA slaběji než R mutant.

Mutanty se naopak od přirozené formy SRSF5 nelišily v lokalizaci a dynamice v buňce. Rozdíly ve vazbě na RNA byly pravděpodobně překryty dalšími faktory, které ovlivňují chování SRSF5, nejspíše protein-proteinovými interakcemi. Mutantní SRSF5 také neměly dominantní efekt na sestřih alternativního EDB exonu fibronektinu.

Získané výsledky poskytují nepřímý důkaz pro hypotézu, že acetylace má vliv na vazbu SRSF5 na RNA. Mutant napodobující acetylovanou formu proteinu se vázal na RNA nejméně, což naznačuje, že acetylace snižuje interakci SRSF5 s RNA. Toto zjištění může být zajímavé i obecněji, jelikož u mnoha dalších proteinů byla také zjištěna acetylace v RNA vazebné doméně, což činí z acetylace nového hráče v regulaci RNA-proteinových interakcí a potažmo i v sestřihu pre-mRNA.

Table of contents

List of abbreviations

AcK	ε-N-lysine acetylation
ASL	argininosuccinate lyase
ATM kinase	ataxia telangiectasia mutated kinase
ATP	adenosine-5′ -triphosphate
BAC	bacterial artificial chromosome
CBP	CREB binding protein
CLIP	crosslinking and immunoprecipitation
CoA	coenzyme A
DAPI	4',6-diamidino-2-phenylindole
DMEM	Dulbecco`s modified Eagle`s medium
DRB	5,6-Dichloro-1-β-D-ribofuranosylbenzimidazole
ECL	enhanced chemiluminiscence
EGFP	enhanced green fluorescent protein
EHHADH	enoyl-CoA, hydratase/3-hydroxyacyl CoA dehydrogenase
FRAP	fluorescence recovery after photobleaching
GAPDH	glyceraldehyde-3-phosphate dehydrogenase
GST	glutathione S-transferase
HAT	histone acetyltransferase
HDAC	Histone deacetylase
HRP	horseradish peroxidase
IGFR	insulin-like growth factor 1 receptor
IPTG	Isopropyl β-D-1-thiogalactopyranoside
iTRAQ	isobaric tags for relative and absolute quantification
MS	mass spectrometry
MDH	malate dehydrogenase
MVH	mouse VASA homolog
NaBu	sodium butyrate
NAD	nicotinamide adenine dinucleotide
NAM	nicotinamide
NLS	nuclear localization signal
PBS	phosphate-buffered saline
PEPCK	phosphoenolpyruvate carboxykinase
PI3K	phosphatidylinositol 3-kinase
PIPES	piperazine-N,N′-bis(2-ethanesulfonic acid)
P-TEFb	positive transcription elongation factor b
qPCR	quantitative polymerase chain reaction
RIP	RNA immunoprecipitation
RRM	RNA recognition motive
RT-PCR	reverse transcription polymerase chain reaction
SAHA	Suberoylanilide hydroxamic acid
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SELEX	systematic evolution of ligands by exponential enrichment
SILAC	stable isotope labeling with amino acids in cell culture
siRNA	small interfering RNA
snRNP	small nuclear ribonucleoprotein
SR proteins	serine/arginine-rich proteins
TBS	Tris-buffered saline
TCA cycle	tricarboxylic acid cycle
TF	transcription factor
TSA	trichostatin A
UTR	untranslated region

Introduction

Lysine acetylation

The acetylation of proteins was discovered almost fifty years ago as a posttranslational modification of histones influencing gene expression (Allfrey et al., 1964) and for a long time it was studied exclusively in this respect. Later on, more targets of acetylation was reported, such as tubulin (L'Hernault and Rosenbaum, 1985) and p53 (Gu and Roeder, 1997), but acetylation was not systematically studied. All these cases fall into the category of ε-N-lysine acetylation (AcK) which I will review in this introduction. I will not focus on structurally similar, but functionally distinct, α-N-terminal acetylation, which is actually the most frequent protein acetylation present in 80% of human proteins and among other functions prevents proteins from entering the secretory pathway, regulates protein-protein interaction and influences their stability (Arnesen, 2011). When people discuss the influence of acetylation on protein function, they often mention that acetylation neutralizes positive charge on lysine residues, but we should not forget that acetylation also changes the overall size of the side chain, its hydrophobicity and potential to form hydrogen bonds.

The view of AcK has changed in last few years and attention is turning from histones to other acetylated proteins as well. For a long time, research on protein acetylation was limited by lack of adequate methods and only individual cases of acetylation were studied. Recently the development in the field of mass spectrometry and first antibodies reactive against acetylated lysine allowed researchers to uncover the true scope of lysine acetylation (Choudhary et al., 2009; Kim et al., 2006). The current view is that acetylation is of equal importance as phosphorylation, regulating many cellular processes, such as gene expression, cell cycle, metabolism, actin cytoskeleton dynamics or aging (Norris et al., 2009). It is widespread and evolutionarily conserved. Many AcK appear in prokaryotes as well, where they regulate mainly metabolism, flagellar motion, translation but also other processes (Jones and O'Connor, 2011).

The research on AcK led to many medically relevant discoveries, for example some inhibitors of lysine deacetylating enzymes are in use or in clinical trials for treatment of cancer, neurodegenerative diseases or as anti-inflammatory drugs. They can be also used for improving the efficiency of induction of pluripotent stem cells (Huangfu et al., 2008). The mechanism of action of lysine deacetylase inhibitors is still unclear, probably different subset of acetylation sites in the cell is influenced in each application.

Histone acetyl transferases (HATs) and histone deacetylases (HDACs)

The acetylation in the cell is maintained in equilibrium by opposing activities of HATs and HDACs. The histone in their name is given by historical circumstances of their discovery. Histones were the first discovered substrate but now we know many non-histone porteins that are targets of HATs and HDACs. Relatively small number of HATs (30) and HDACs (18) regulate thousands of acetylation sites, which makes AcK different from phosphorylation that is carried out by 500 kinases and about 150 phosphatases. Similarly the ubiquitination is carried out by 1000 E3 ligases and 80 deubiquitinylating enzymes (Guan and Xiong, 2011). So what is the determinant of acetylation specificity? The sequence around the acetylation sites, though being somewhat conserved, cannot itself explain acetylation specificity. The major determinant of AcK is still matter of debate, protein-protein interactions definitely play an important role.

HATs occur usually in macromolecular complexes and we can divide them into three families CBP/p300, MYST and GNAT. They have multiple subunits and the function of the catalytically active subunit is regulated in terms of activity and substrate specificity by composition of the whole complex. For this reason, *in vitro* experiments with purified HATs often give misleading results. HAT complexes contain many chromatin-binding domains such as bromodomain that binds acetylated lysines and chromodomain, Tudor domain, WD40 repeats and PHD fingers that all bind various methylated lysines. HATs use acetyl-CoA as a source of acetyl group, which links their activity to the metabolic state of the cell.

CBP/p300 family of proteins were discovered as coactivators of transcription, that are recruited by transcription factors (TFs) to transcription units. The acetyltransferase activity, which is at the core of their coactivator function, was discovered later. Currently there are over 40 TFs known to interact with p300 or CBP (CREB binding protein), notably p53, nuclear hormone receptors, Fos, Jun, STAT or Smad (Chen and Li, 2011). CBP and p300 are two highly homologous proteins with mostly overlapping functions but a few specific ones as well (Kalkhoven, 2004). These two proteins are large and possess various domains like HAT, Bromodomain and several transcription factor-interacting domains KIX, C/H1, 2 and 3. The pivotal substrates of p300 are histones and transcription factors. Besides its enzymatic function p300 serves as a scaffold for transcription factors or as a bridge for TFs with the general transcription machinery (Chen and Li, 2011). p300 is activated by autoacetylation and also by phosphorylation mediated by cyclin-dependent kinases or by Akt.

Phosphorylation in general increases resistance of p300 to proteasomal degradation. The level of p300 in the cell is often crucial for transcription regulation, because individual transcription units compete for a limited pool of p300. Activating role of p300 is also regulated by its export to cytoplasm, where it has another unrelated function as a p53 E4 ubiquitin ligase. It has a ubiquitin ligase domain on N-terminus that is active only in cytoplasm. This destabilizing effect on p53 in the cytoplasm is interesting, because acetylation of p53 in the nucleus has the opposite effect and stabilizes p53 (Chen and Li, 2011). p300 participates also in other processes that require chromatin remodeling, such as DNA replication and DNA repair.

CBP and p300 are essential and mutations or other inactivation of CBP gene that lead to reduction of active protein by 50 % cause Rubinstein-Taybi syndrome (RTS), which manifests itself as developmental abnormalities, mental retardation and increased incidence of cancers. The mutations causing RTS are mostly found in the acetyltransferase domain, but a mutation in bromodomain was also reported (Kalkhoven, 2004). The role of CBP and p300 in cancer is not resolved, in some cases it behaves as tumor suppressor. It is mutated in some cancers of epithelial origin but larger studies of tumors and cancer cell lines failed to discover frequent mutations in CBP. CBP and p300 play a distinct role in acute myeloid leukemia. Often there appear translocations causing fusion of p300 or CBP with MLL or MOZ proteins producing a protein with altered HAT activity (Kalkhoven, 2004).

The HATs from MYST family are phylogenetically widespread predominantly nuclear proteins that regulate most processes localized to this cellular compartment. The most studied member of the family, Tip60 acetyltransferase, is an important player in transcription regulation. It acetylates both histones and transcription factors such as p53 or c-Myc. Tip60 also promotes DNA damage response to double-strand breaks by acetylating histones in the site of break and by activating the ATM kinase (Sapountzi and Cote, 2011).

The GNAT family contains Gcn5, the first HAT discovered, PCAF (p300/CBP-associated factor) and many related proteins. They function similarly to CBP/p300 as transcription coactivators acetylating histones and nonhistone proteins. Gcn5 is essential for embryonic development in mice and Drosophila unlike PCAF. PCAF knockout mice develop normally but have elevated levels of Gcn5, which suggests that Gcn5 is able to replace PCAF in most of its functions. Gcn5 like other HATs participates in several well known macromolecular complexes. In yeast it is the large SAGA complex and smaller ADA complex, in humans its

SAGA homologs STAGA and TFTC and ADA homolog ATAC. Similar complexes were found in Drosophila as well. Besides the HAT domain, PCAF contains also E3 ubiquitin ligase domain and ubiquitinates p53 (Nagy and Tora, 2007).

In mammals there are fivevclasses of HDACs with total 18 members, which rises interesting questions about their specificity and redundancy. Class I, IIa, IIb and IV are Zn^{2+} dependent and class III, sirtuins, are NAD^{+} dependent. Class I HDACs are expressed ubiquitously and they are components of various corepressor complexes in the nucleus such as NuRD or Polycomb repressive complexes so their major substrates are acetylated histone N-termini and their role is silencing gene expression (there are also reported cases of gene activation (Haberland et al., 2009)). Mouse knockouts of members of all classes, except for class IIa, are often lethal embryonically or around birth suggesting important functions for HDACs in development (Haberland et al., 2009). The severe phenotype might also be caused by disintegration of complexes HDACs would normally be found in.

Class I HDACs are the usual target of HDAC inhibitors, unlike class IIa HDACs that are resistant to most inhibitors in use (Delcuve et al., 2012). A recent study (Bantscheff et al., 2011) put most of the data about specificity of HDAC inhibitors in question, because instead of using purified HDACs they examined HDAC inhibition within the macromolecualr complexes where they naturally occur. The inhibition of particular HDAC proved to be dependent not only on the class of HDAC but also on the complex it was part of. There is a striking difference between mostly lethal HDAC knockouts and mild effect of chemical HDAC inhibition. It probably arises from disruption of protein complexes containing HDACs that are not properly formed in the knockouts (Haberland et al., 2009). Most inhibitors in use are broad-spectrum and affect all but class IIa HDACs and sirtuins. Global inhibition of HDACs has often surprisingly specific and beneficial effect in treatment of many diseases (**Table 1**). The utility of HDAC inhibitors against so many miscellaneous diseases reflects the ubiquitous cellular regulation by protein acetylation. First at the level of transcription and also in regulation of function of non-histone proteins (Haberland et al., 2009).

Disease	Proposed mechanism
Humans	
HIV infection	De-silencing of latent virus
Cutaneous T-cell lymphoma	Upregulation of tumor-suppressor genes, induction of apoptosis
GPI deficiency	Increased PIGM expression by hyper-acetylation of histones on promoter
Ulcerative colitis	Inhibition of NF-κB activation
Sickle cell disease	Increase in fetal haemoglobin expression
Mice	
Arthritis	Inhibition of TNFα expression and of inflammation
Asthma	Inhibition of cytokine expression and T-cell infiltration
Autoimmune encephalitis	Upregulation of antioxidant, antiexcitotoxicity and proneuronal factors
Colitis	Suppression of pro-inflammatory cytokines
Cardiac hypertrophy	Unknown
Dementia	Dendrite sprouting, increased synapse number
Graft versus host disease	Reduction of pro-inflammatory cytokines
Hepatitis	Inhibition of TNFα and INFγ
Muscular dystrophy	Induction of follistatin
Systemic lupus erythematosus	Downregulation of pro-inflammatory cytokines
Spinal muscular atrophy	Activation of survival motor neuron 2 gene
Rats	
Haemorrhagic shock	Reduction of TNFα expression
Brain trauma	Inhibiting neuroinflammation

Table 1: Diseases treated with HDAC inhibitors and their proposed mechanism of action (Haberland et al., 2009).

Class II HDACs are both nuclear and cytoplasmic and their export from the nucleus is activated by phosphorylation and following 14-3-3 protein binding. Class IIa proteins are expressed tissue-specifically and they are not themselves catalytically active, but they form a heterodimer with class I HDACs. Knockout mice for class IIa HDACs are mostly viable, probably due to redundancy among them, even though the pattern of their expression is diverse. Class IIb HDACs target other proteins rather than histones mostly in the cytoplasm and cytoplasmatic membrane. One member of the family, HDAC6, is the deacetylase of α-tubulin and of diverse proteins regulating the actin cytoskeleton. So HDAC6 is a potent regulator of cell motility, which makes it an interesting anticancer drug target. Another well known target of HDAC6 is HSP90. HDAC6 is also associated with aggresome, a proteinaceous inclusion body that forms when a cell is overwhelmed by misfolded proteins implicating a role for HDAC6 in neurodegenerative disease. Mouse knockout of HDAC6 has surprisingly

mild phenotype, only highly acetylated microtubules, which probably stems from the redundancy with the other member of class IIb, HDAC10 (Delcuve et al., 2012). The only class IV HDAC, HDAC11, is not very well explored.

Sirtuins (class III HDACs) are a family of proteins conserved from prokaryotes to mammals. They were first discovered in yeast as regulators of heterochromatin silencing, but they gained attention because they also prolonged yeast lifespan (Kaeberlein et al., 1999). Soon their main function as deacetylases was discovered (Imai et al., 2000). Unlike other HDACs that simply hydrolyze acetylated lysine, sirtuins use NAD^+ (not NADH or $NADP^+$) in the deacetylation reaction. They cleave it yielding lysine, nicotinamide and 2′-O-acetyl-ADP-ribose. The product of this reaction, nicotinamide, is an inhibitor of sirtuin activity. Sirtuins are dependent on the cellular levels of NAD^+ or NAD^+/NADH ratio, which predestines them to be sensors of cellular metabolic state (Houtkooper et al., 2012). Sirtuins have also other enzymatic functions related to acetylation such as demalonylation and desuccinylation and ADP-ribosylation, which arises from the NAD^+ cleavage.

The best explored are nuclear sirtuin SIRT1 and mitochondrial SIRT3. From studies on transgenic mice it emerged that SIRT1 is one of the factors that mediate the benefitial effects of calorie restriction regimen, such as reduction of age-related health problems. SIRT1 improves the sensitivity of body to insulin and also insulin secretion in pancreas by deacetylating histones and other proteins like transcription factors. It is a promising drug target for treatment of type-2 diabetes (Houtkooper et al., 2012). SIRT3 is the main mitochondrial deacetylase regulating many mitochondrial metabolic enzymes. It is also important during calorie restriction and other stresses, where it reduces ROS production by activating SOD2 and through other mechanisms.

A lively area of research is the role of sirtuins in longevity. As calorie restriction regimen increases longevity in wide spectrum of species, it was speculated that sirtuins as one of the enzymes activated by calorie restriction could themselves be longevity factors. But besides yeast, the role of sirtuins in longevity shown in the initial studies is now questioned by newer data from *Caenorhabditis elegans*, *Drosophila melanogaster* and mice overexpressing SIRT1 showing none or negligible extension of lifespan (Houtkooper et al., 2012). But other model emerges. Sirtuins do not extend lifespan, but rather healthspan by their favourable effect on age-related metabolic diseases and in situations of metabolic stress like high fat diet demonstrated in SIRT1 knockin mice (Houtkooper et al., 2012). Most of the studies in this

field were done on SIRT1 but other sirtuins, which are the target of intensive research now, might still regulate longevity.

Activation of benefitial effects of sirtuins by rather uncomfortable calorie restriction led to the thought that instead, small molecule agonists of SIRT1 could be used to improve the age-related or unhealthy diet-triggered metabolic complications in humans. The first discovered SIRT1 activator was a plant polyphenolic compound resveratrol, which mimicked the effect of calorie restriction and extended the lifespan of yeasts (Howitz et al., 2003). Resveratrol activates SIRT1 in mice and humans as well. Although it turned out, that SIRT1 is probably not a direct target of resveratrol. It enhances mitochondrial function and protects mice againts diet-induced obesity. It also extends their lifespan, when they are metabolically stressed by high fat diet. Artificial agonists of SIRT1 were also prepared, the most potent one, SRT1720, has similarly beneficial effects as resveratrol (Houtkooper et al., 2012).

Lessons learned from global acetylation studies

Kim and colleagues published the first global acetylation study in 2006 where they identified 388 acetylation sites in 195 proteins (Kim et al., 2006) and by which they more than doubled the number of acetylated proteins known at that time. This study was technically feasible because they used a newly prepared anti-acetylated lysine antibody for initial enrichment of trypsin-digested peptides. Some kind of enrichment is necessary in such a study and is a common approach also in phosphoproteomics. As the material for analysis they used HeLa cells and mouse mitochondria. Their most important finding was probably the unexpected discovery of widespread AcK in mitochondria, which indicates a connection of acetylation to energy metabolism. I`ll discuss that in a separate section later. Among the acetylated proteins were also chaperones, RNA splicing and translation factors and many regulators of actin cytoskeleton and also actin itself.

They tested the role of acetylation on actin cytoskeleton by agonists and antagonists of histone deacetylases and found that agonists (causing low acetylation) destabilized and antagonists (causing high acetylation) stabilized filamentous actin, which resulted in formation of more stress fibers. This is in concordance with other observation about acetylation-mimicking mutation (lysine to glutamine) in regulator of Rho GTPases, RhoGDI, that caused stress fiber formation unlike lysine mutation to arginine, which mimics nonacetylated lysine and did not significantly change actin cytoskeleton.

The number of identified acetylation sites allowed for the first search of acetylation motive in the surrounding protein sequence. It was different in nuclear and cytoplasmic or mitochondrial proteins, which probably reflects distinct acetylating enzymes acting in different cellular compartments. As for the secondary structures, AcK is found more often in structured regions (helices) compared to nonacetylated lysines and less frequently in unstructured regions than the nonacetylated lysines (histones being an exception). This makes them different from phosphorylations, which correlate with unstructured regions.

The next study confirmed the findings of Kim and colleagues and broadened them significantly (Choudhary et al., 2009). After the enrichment for acetylated peptides they employed further fractionation which helped to overcome the problem that the sample after affinity step is dominated by a few most abundant acetylated peptides. The list of acetylation sites was expanded to 3600 in 1750 proteins from three different human cancer cell lines. The acetylomes of these three cell lines significantly overlapped indicating that the study got close to mapping the whole human acetylome. Though more acetylations might still be found under different growth conditions and in different cell lines or tissues. Authors gained many important insights from global analysis of the human acetylome. The amount of data collected in this study also allows for computational prediction of acetylation sites (Wang et al., 2012).

The acetylated proteins were found to be more evolutionarily conserved when compared to the whole proteome. Acetylation was overrepresented in several Pfam protein domains. Specifically nuclear domains, such as RNA binding motives and helicases. On the other hand the most underrepresented domains were membrane seven helix receptors, EGF-like and peptidase domains. AcK was also significantly enriched in proteins that are part of big protein complexes such as HAT complexes or chromatin-remodeling complexes. Further analysis unexpectedly revealed clusters of closely interacting acetylated proteins taking part in various cellular processes such as RNA splicing, DNA repair, cell cycle, cytoskeleton reorganization, protein folding, nuclear transport or within the ribosome. The high acetylation of RNA binding motives and overall high acetylation of splicing proteins reported in this study were the primary motivation for starting my project.

They also quantitatively analyzed (by SILAC) global effect of two HDAC inhibitors. SAHA, a broad spectrum inhibitor of HDACs (except for sirtuins) and MS-275, which is specific for class I HDACs. Both inhibitors were surprisingly specific and increased acetylation more than

two fold at only 10% of acetylation sites. SAHA was more potent inducer of lysine acetylation. As expected, the mitochondrial acetylation was not influenced. The set of influenced sites by these two inhibitors overlapped (differently for particular cell line) at around 50%. At some well known acetylated proteins each inhibitor had distinct effect. For example SAHA upregulated acetylation of HSP90 unlike MS-275 and MS-275 increased acetylation of p53 and SAHA had no effect. Such specificity of inhibitors, which have also different mechanism of action, explains their distinct effects in research and in clinical applications.

Acetylation sites or events can be divided into two groups. In the first category, there are proteins with only one or very few exceptionally specific acetylation sites, but they are sufficient for altering the protein properties and serve as a switch. Secondly, some proteins have a patch of more lysine residues that can be all acetylated with the same outcome. The acetylations are interchangeable, what matters is the overall charge of that protein segment. It was noticed already before the study of Choudhary and colleagues was published (Yang and Seto, 2008).

Other study focused on acetylation of proteins in mitochondria and cytoplasm in human liver (Zhao et al., 2010). They described 703 previously unknown acetylated proteins. Their dataset included 70% of acetylated proteins discovered by Kim and colleagues in mouse liver but only 14% of acetylated proteins from human cancer cells reported by Choudhary and collegues. It indicates that acetylation is conserved in the same cell type across species but various cell types in one body can have completely different acetylation patterns. They focused on metabolic enzymes because they found acetylation in virtually every enzyme from glycolysis, TCA cycle, gluconeogenesis, glycogen and fatty acid metabolism or urea cycle.

They mapped in detail the influence of nutrients and HDAC inhibitors on acetylation of several selected enzymes and how acetylation affects their activity. They used Chang human liver cells expressing FLAG-tagged proteins for immunoprecipitation, and iTRAQ isobaric tags for MS quantification of acetylated peptides. The EHHADH enzyme from fatty acid metabolism was positively regulated by acetylation. Acetylation of EHHADH increased after treatment with HDAC inhibitors nicotinamide and TSA from 40% to 70% and activity of the enzyme doubled. Activity and acetylation of EHHADH was also inceased by additon of fatty acids to the culture medium implicating acetylation in physiological regulation of EHHADH

function. The acetylation of TCA cycle enzyme malate dehydrogenase (MDH) increased from 30% to 70% after treatment of cells with nicotinamide and TSA and activity of the enzyme doubled. Incubation of cells in high glucose conditions increased again the acetylation and activity of MDH. Consistently *in vitro* deacetylation of MDH decreased its activity and MDH mutant with acetylated lysines changed to arginines did not respond to either HDAC inhibitors or increased glucose by increased enzymatic activity.

The urea cycle enzyme, argininosuccinate lyase (ASL), was negatively regulated by acetylation. The HDAC inhibitors increased its acetylation, which lowered ASL enzymatic activity. Growing cells in high concentration of amino acids decreased the acetylation and increased activity of ASL. And *in vitro* deacetylation of ASL increased its activity. The urea cycle is connected to TCA cycle and gluconeogenesis because in low glucose conditions, the cells use fumarate from the urea cycle as an energy source and for gluconeogenesis. So in low glucose conditions urea cycle enzymes are more active than in high glucose conditions and indeed in high glucose conditions ASL was more acetylated and less active. It shows that acetylation can be a signal, which connects more metabolic pathways together.

In the last case it was shown that acetylation destabilizes gluconeogenic enzyme PEP carboxykinase (PEPCK) in high glucose conditions, when gluconeogenesis is not required. In high glucose conditions PEPCK was more acetylated, but the overall amount of protein dropped. Treatment with HDAC inhibitors had similar effect. On the contrary PEPCK was stable in glucose-free medium. The PEPCK mutant with acetylated lysines swapped to arginines was also more stable than the wild type.

All these results show that a substantial fraction of an enzyme in the cell can be acetylated and this acetylation changes dynamically in response to nutrient availability and metabolic condition of cells. The signal relayed by acetylation coordinates different metabolic pathways in response to changing conditions. I will expand these ideas in the next section.

Regulation of cellular metabolism by acetylation

Acetylation is intimately linked to energy metabolism because the donor of acetyl group in the reaction is acetyl-CoA. Deacetylation by sirtuins as already mentioned requires oxidized NAD. In the most general view, we can say that acetylation dominates when

nutrients are abundant, acetyl-CoA is high, NAD^+ level is low thus SIRT activity is low, which leads to high acetylation. During starvation it is vice versa and deacetylation steps in.

AcK is not the only posttranslational modification connected with metabolism. Basically all PTMs are generated from metabolites and therefore might be sensitive to nutrient levels. There is evidence supporting this view at least for glycosylation and acetylation (Wellen and Thompson, 2012). This is likely also for modifications analogous to acetylation like butyrylation, propionylation and succinylation, which were found in past years (Garrity et al., 2007; Zhang et al., 2009; Zhang et al., 2011). They all have in common that they can be found in cells as CoA derivatives.

AcK is involved in regulation of metabolism in all domains of life. Besides the mouse and human examples (Choudhary et al., 2009; Kim et al., 2006; Zhao et al., 2010) there were reports describing frequent acetylation of metabolic enzymes in *Escherichia coli* (Yu et al., 2008; Zhang et al., 2009) and *Salmonella enterica* (Wang et al., 2010) and there are few reported acetylated proteins in Archaea as well (Jones and O'Connor, 2011), probably many more would be found if someone performed a comprehensive survey. Yu and colleagues discovered 85 acetylated proteins in *E.coli* most of which were metabolic enzymes or proteins involved in proteosynthesis. The majority of proteins were acetylated in stationary phase and deacetylated in exponential phase suggesting a regulatory role for acetylation in bacterial physiology (Yu et al., 2008).

The second study on E. coli (Zhang et al., 2009) acquired AcK dataset of similar size but not much overlapping with the previous study, which probably reflects undersampling and difference in bacterial strains and growth conditions between these two studies. Zhang and colleagues came to similar conclusions and furthermore they found also increased AcK in stress conditions and in stress response proteins. They also highlighted that lots of acetylations are conserved from prokaryotes to mammals (Zhang et al., 2009).

The study on *Salmonella* revealed 191 acetylated proteins out of which one half were metabolic enzymes of assorted pathways. 90% of enzymes in basic metabolic pathways like glycolysis and TCA cycle were acetylated (Wang et al., 2010). They grew cells in fermenting conditions with glucose as carbon source (high glycolysis) or in oxidative conditions with citrate (high gluconeogenesis). Higher AcK was found in bacteria growing on glucose and it was demonstrated that acetylation stimulates glycolysis and inhibits gluconeogenesis. The authors exploited an advantage of *Salmonella* having only one major HAT and HDAC (NAD^+

dependent) so deletion mutants with much reduced HAT and HDAC activity could be easily created. They prepared those mutants and as expected the deacetylase mutant had more acetylated proteins and the acetyltransferase mutant had lower overall acetylation. When they compared the growth rate of the deacetylase mutant on glucose with the wild type, it grew faster. On citrate it grew slower. The experiments with acetyltransferase mutant gave the opposite results. It is in agreement with the model that acetylation stimulates glycolysis and inhibits gluconeogenesis.

This hypothesis was supported also by close examination of GAPDH, an enzyme of both glycolysis and gluconeogenesis. When this enzyme was acetylated, it favoured the forward glycolytic reaction and when it was deacetylated, it was more effective at the reverse gluconeogenic reaction (Wang et al., 2010). These results were also confirmed by feeding the bacteria with ^{13}C glucose or citrate and measurements of metabolite levels *in vivo*. For example when grown on glucose, the deacetylase mutant had two times bigger ratio of glycolysis/gluconeogenesis intermediates compared to the acetyltransferase mutant (Wang et al., 2010). The major contribution of this study is characterizing how AcK regulates activities of metabolic enzymes and how it influences flow of carbon through metabolic pathways. Most of the findings might be valid for eukaryotes as well, e.g., human GAPDH is acetylated as well. It remains to be tested if the acetylation influences the enzyme the same way as in *Salmonella*.

The requirement of acetylation for glycolysis could be potentially used against various types of cancers, because they are dependent mostly on glycolysis and subsequent fermentintation to lactate. Their rate of glycolysis is hundred times higher than that of healthy tissues (Kim and Dang, 2006), so for example HAT inhibitors or sirtuin agonists could considerably inhibit tumor growth.

The situation of acetylation in mitochondria is more ambiguous. Elevated fatty acid oxidation in low nutrient situations would increase acetyl-CoA and therefore increase acetylation, on the other hand nutrient shortage activates mitochondrial SIRT3, which deacetylates and activates oxidative enzymes and scavengers of reactive oxygen species (Guarente, 2011). Activation of these enzymes accounts for the therapeutic effect of sirtuins on age-related diseases. The change in mitochondrial acetylation during calorie restriction is tissue-specific as was shown in mice (Schwer et al., 2009). In liver mitochondria, AcK was increased on 72 enzymes. In brown fat calorie restriction had the opposite effect and

decreased mitochondrial acetylation. In other tissues like brain, kidney or heart AcK did not change (Schwer et al., 2009). It is not surprising for a multicellular organism, because the liver and fat are professional metabolizing tissues that sense, store and metabolize nutrients to provide stable supply of glucose and lipids into the bloodstream for the rest of the body.

Further evidence for close relationship of acetylation and metabolism, apart from the metabolic enzymes, is the regulatory role of acetylation in insulin-like growth factor receptor signaling. Acetylation of PTEN inhibits its phosphatase activity, therefore enhancing the PI3K signaling. Also the downstream targets of IGFR pathway, the FoxO transcription factors, are regulated by AcK via several mechanisms. It reduces their nuclear import, increases their ubiquitination and attenuates their binding to DNA (Yang and Seto, 2008).

Mechanisms of influence of acetylation on protein function

Here I would like to summarize various ways how AcK can influence protein function, which I will document with some representative examples (they are not ment as an exhaustive list). A few years old review (Spange et al., 2009) contains a thorough list of acetylated proteins known at that time and the biological implications of their acetylation (**Table 2**). Since then our knowledge of acetylated proteins exploded, which made compilation of the comprehensive lists of acetylated proteins hard or even impossible to create.

Biological implication	Proteins affected by acetylation	
Protein stability	Acetylation increases stability p53, p73, Smad7, c-Myc, Runx3, AR, H2A.z, E2F1, NF-E4, ER81, SREBP1a, HNF6, BACE1	Acetylation decreases stability GATA1, HIF-1α, pRb, SV40 T-Ag
DNA binding	Increased DNA binding p53, SRY, STAT3, GATA transcription factors, E2F1, p50 (NFκB), Erα, p65 (NFκB), c-Myb, MyoD, HNF-4, AML1, BETA2, NF-E2, KLF13, TAL1/SCL, TAF(I)68, AP endonuclease	Decreased DNA binding YY1, HMG-A1, HMG-N2, p65 (NFκB), DEK, KLF13, Fen-1
Gene expression	Transcriptional activation p53, HMG-A1, STAT3, AR, ERα (basal), GATA transcription factors, EKLF, MyoD, E2F1, p65(NFκB), GR, p73, PGC1α, MEF2D, GCMa, PLAG1, PLAGL2, Bcl-6, β-Catenin, KLF5, Sp1, BETA2, Cart1, RIP140, TAF(I)68	Transcriptional inactivation Erα (ligand-bound), HIF-1α, STAT1, FOXO1, FOXO4, RIP140
Protein interactions	Enhanced STAT3, AR, EKLF, Importin A, STAT1, TFIIB, α-Tubulin, actin, cortactin	Decreased p65(RelA), Ku70, Hsp90
Localisation	Ac → nucleus PCAF, SRY, CtBP2, POP-1, HNF-4, PCNA Sub-nuclear WRN, PCNA	Ac → cytosol c-Abl, p300, PAP
mRNA stability	Increased p21, Brm	Decreased Tyrosinhydrolase (Th), eNOS
Enzymatic activity	Enhanced p300, ATM	Decreased PTEN, HDAC1, Mdm2, ACS, Neil2, Polβ
Mitochondrial proteins	ACS (Ac-CoA-Synthetase), Sod1/2, Profilin I, Thioredoxin; multiple components of metabolic and oxidative phosphorylation machinery	
Viral proteins	E1A, S-HDAg, L-HDAg, HIV Tat, SV40 T-Ag	

Table 2: Various implications of protein acetylation (Spange et al., 2009).

AcK can modulate enzymatic activity, which could be documented on p300 (Thompson et al., 2004) that is activated by autoacetylation just like kinases and phosphatases are regulated by phosphorylation. Such autoacetylation was discovered early on in *in vitro* assays, where the acetyltransferase acquired the radioactive acetyl just like the assayed protein. ATM kinase is activated by acetylation in response to DNA damage (Sun et al., 2005). Acetylation of lysine that is important for ATP binding inhibits various cyclin-dependent kinases (Choudhary et al., 2009). Acetylation serves as an on/off switch in many metabolic enzymes such as acetyl–coenzyme synthetase, which is switched on by Sir2 mediated deacetylation (Starai et al., 2002) or all the metabolic enzymes described earlier (Zhao et al., 2010). This mechanism of regulation of acetyl–coenzyme synthetase is conserved from prokaryotes to mammals (Guan and Xiong, 2011).

Acetylation can change subcellular localisation of proteins by covering the nuclear localisation signal (NLS). For example Inhibitor of Growth 4 (ING4) is acetylated at three lysines within its NLS (Kim et al., 2006). The splicing factor DEK relocates after acetylation from nucleoplasm to nuclear speckles (Cleary et al., 2005), which are nuclear storage compartment for RNA processing and transcription factors so acetylation likely inhibits DEK function in splicing.

AcK crosstalks with other posttranslational modifications. Very well documented is the interplay of various modifications on N-terminal tails of histones known as the histone code

(Jenuwein and Allis, 2001). During gene silencing, acetylation is replaced by methylation on several lysines in histones, e.g. H3K9, H3K27, H4K20 (Kouzarides, 2007). HATs can be recruited to chromatin by chromodomains recognizing H3K4 and H3K36 methylation. H3K9 and H3K27 acetylation is also enhanced by H3S10 and H3S28 phosphorylation, which represents a way how to relay a transient signalization through phosphorylation into lasting cellular response based on acetylation. Such acetylated and phosphorylated KS motive is present in more proteins (Yang and Seto, 2008). ADP-ribosylation can also crosstalk with the previously mentioned modifications. Their relationship can be antagonistic in the sense that they are competing for the same lysine residue, but also agonistic, because chromatin opening by acetylation promotes ADP-ribosylation (Hottiger, 2011).

Several modifications interplay during p53 regulation, which is known in great detail. AcK in C-terminal domain by p300 is important for p53 stabilisation, because the same lysines are also targets of ubiquitination. But prior phosphorylation of N-terminal part is required for p300 binding to p53 (Yang and Seto, 2008). Other proof of connection of ubiquitination and acetylation is the widespread acetylation among nuclear ubiquitin ligases and deubiquitinating enzymes (Choudhary et al., 2009). Sumoylation of K524 of Ran GTPase1 is required for nuclear import, the same lysine residue can be also acetylated (Choudhary et al., 2009; Mahajan et al., 1997). Acetylation at four lysines decreases binding of 14-3-3 proteins to their substrates (mostly serine or threonine-phosphorylated proteins), which was established by acetylation-mimicking mutations (lysine to glutamine mutation) in 14-3-3 (Choudhary et al., 2009), which is an example of crosstalk of acetylation with phosphorylation and also documents other typical function of acetylation in altering protein-protein interactions.

Prime example are proteins containing a bromodomain, which is an acetylated lysine-binding domain. It is typical for nuclear proteins involved in chromatin remodelling or transcription control, in cytoplasm or mitochondria such a domain has not been discovered. But there are many specific examples, e.g., acetylation within substrate binding domain of HSP90 abolishes its binding to the substrates (Choudhary et al., 2009; Scroggins et al., 2007). The interferon signalling is activated by acetylation, which promotes the interaction between the receptor and IFN regulatory factor 9 (IRF9) and also downstream between activated STAT proteins (Tang et al., 2007). Somewhat special example is acetylation of α-tubulin, which is widespread from unicellular eukaryotes to vertebrates. It was observed

that stable microtubules are more acetylated than the dynamic ones and HDACs induce microtubule depolymerization. Acetylation also triggers the formation of a docking site for motor proteins (both kinesins and dyneins) on microtubules (Yang and Seto, 2008).

Acetylation alters binding of proteins to DNA. Tip60-mediated acetylation in DNA-binding domain of p53 regulates, whether it will activate cell-cycle arrest or apoptotic cellular program (Sykes et al., 2006; Tang et al., 2006). The binding of nuclear hormone receptors to DNA is also regulated by acetylation, in most cases acetylation improves their binding capability (Wang et al., 2011). Similar observation was made for GATA transcription factors (Boyes et al., 1998) E2F transcription factors (Martinez-Balbas et al., 2000) or MyoD (Polesskaya et al., 2000).

Acetylation alters RNA-protein interaction

Last but not least acetylation alters RNA-protein interaction, which I will analyze in more detail. AcK neutralizes the positive charge on this residue and reduces the electrostatic interactions of negatively charged RNA with positively charged RNA-binding domain. The role of electrostatic interaction in RNA binding is well studied (Clery et al., 2011; Dominguez and Allain, 2006; Garcia-Garcia and Draper, 2003; GuhaThakurta and Draper, 2000; Law et al., 2006). It is important in the initial recognition phase of RNA and protein at longer distances up to 11 Å then short distance interactions gain more importance (the 'lure and lock' model (Law et al., 2006)). Electrostatic interactions do not have to be the exclusive mechanism. As stated earlier acetylation changes also size, hydophobicity and potential to form hydrogen bonds of the lysine side chain. Such alteration can trigger conformational change of the whole RNA-binding domain.

The first example comes from HIV biology. The HIV protein Tat is essential for HIV transcription elongation. It binds a stem-loop structure at the 5′ end of HIV transcripts and recruits cyclin T1, part of P-TEFb complex (together with Cdk9) that phosphorylates C-terminal domain of RNA polymerase II and triggers transcription elongation (Kiernan et al., 1999). Tat is acetylated on two lysines with different outcomes. Acetylation on lysine 28 by PCAF enhances interaction with P-TEFb and lysine 50 acetylation weakens Tat interaction with HIV RNA, which enables Tat recycling. Recycling of Tat is important, because the expression of this protein is very low. HIV tanscription is also boosted by HDAC inhibitor TSA (Kiernan et al., 1999). Later it was shown that K28 is also important for RNA binding, because

it stabilizes the Tat-RNA-cyclin T1 ternary complex (D'Orso and Frankel, 2009). This is counteracted by K28 deacetylation by HDAC6 (Huo et al., 2011).

The role of acetylation in RNA binding was also described in 2004 for Sam68 protein (Babic et al., 2004). It has diverse roles in cell signaling and regulation of cell cycle or pre-mRNA splicing. It was shown that increased acetylation of Sam68 enhances its interaction with RNA. Acetylation of Sam68 was highest in tumorigenic breast cancer cell line implicating enhanced RNA binding of Sam68 in cancer cell proliferation. It was shown later that Sam68 is acetylated directly in its RNA-binding domain (Choudhary et al., 2009). Acetylation enhancing RNA binding is somewhat counterintuitive. Most likely acetylation enhances other interactions which compensate for the loss of electrostatic interaction of lysine side chain with phosphates in RNA.

The opposite situation, where acetylation weakens RNA-protein interactions was documented for mouse VASA homolog (MVH or DDX4) in regulation of mRNA granule called chromatoid body (CB) in mouse spermatogenesis (Nagamori et al., 2011). Interestingly, the effect of acetylation on RNA binding was mRNA specific, some mRNAs remained bound even after MVH acetylation. MVH was localized to CB and bound approximately 800 mRNAs and triggered their translational arrest. A significant group among them were testis- or spermatogenesis-specific transcripts or translation regulators like eIF4B. Upon acetylation MVH released mRNA of eIF4B, which was involved in subsequent translational activation. Acetylation of MVH was developmentally regulated. The MVH acetyltransferase Hat1 associated with CB only in certain developmental stage. Acetylation by Hat1 transmits here the endocrine or metabolic signals to the developing sperms (Nagamori et al., 2011).

Reversible AcK is employed also in spreading of the dosage compensation complex around the X chromosome in *Drosophila* males (Buscaino et al., 2003). Dosage compensation in males upregulates expression from the single X chromosome to match the expression from two female X chromosomes. The MSL-3 protein is part of the dosage compensation complex, where it interacts with roX2 non-coding RNA. Upon acetylation of MSL-3 by MOF the interaction with roX2 is lost and MSL-3 dissociates and the complex falls apart. MSL-3 is then rapidly deacetylated, which allows the complex to reassociate at a nearby chromatin site (Buscaino et al., 2003). MSL-3 binds RNA with a chromodomain and the single acetylated lysine is not directly in the domain (lies nearby), which raises the question about the mechanism how acetylation interferes with RNA binding in this case.

Similar mechanism as in *Drosophila* is applied in *de novo* silencing of rDNA by NoRC chromatin remodeling complex (Zhou et al., 2009). The nonacetylated TIP5 subunit of NoRC is recruited to rDNA promoter by interaction with promoter RNA (pRNA; non-coding RNA transcribed from the rDNA promoter). Then TIP5 is acetylated by MOF and TIP5 dissociates from pRNA which triggers the displacement of the promoter-bound nucleosome. Later SIRT1 deacetylation allows TIP5 to reassociate with pRNA and histones around the promoter are modified with heterochromatin marks. The involvement of SIRT1 provides a link to metabolic state of the cell so that appropriate amount of rRNA is produced (Zhou et al., 2009).

Other process where RNA-protein interactions are probably influenced by acetylation is pre-mRNA splicing. Inhibitors of both HATs and HDACs inihibit splicing *in vitro* (Kuhn et al., 2009) and each inhibitor arrests the splicing reaction at specific stage. Though the protein composition of stalled spliceosomes has been establihed, the particular acetylated protein responsible for observed splicing defects remains to be identified (Kuhn et al., 2009).

Acetylation influences RNA-protein interactions in diverse cellular processes. In my thesis I was testing whether it is significant also for the SR protein family of splicing regulators.

SR proteins

The serine/arginine-rich proteins are important regulators of both constitutive and alternative splicing (Long and Caceres, 2009; Shepard and Hertel, 2009). They regulate splicing of both major (U2-dependent) and minor introns (U12-dependent). Canonical SR proteins contain one or two RNA binding motives (RRM) on N-terminus followed by repeats of arginine-serine dipeptides called the RS domain. It must be at least 50 amino acids long with more than 40% of amino acids being serines and arginines (Manley and Krainer, 2010). This definition is rather arbitrary, because there is 51 proteins in humans containing both RRM and SR domain, but not in this order and not all of them function in splicing, which leaves us with 12 true SR proteins (Manley and Krainer, 2010). Other authors calculated slightly different numbers (**Table 3**). Older definitions of SR proteins, now abandoned, included their recognition by m104 monoclonal antibody or their ability to complement splicing-deficient HeLa cytoplasmic S100 extract. By that definition there were 7 classic SR proteins in humans, ASF/SF2, SC35, SRp20, SRp75, SRp40, SRp55 and 9G8 now renamed as SRSF1 to SRSF7 (Manley and Krainer, 2010).

Organism	#SRs
Glycine max	25*
Populus trichocarpa	20
Arabidopsis thaliana	18
Vitis vinifera	9
Zea mays	22
Sorghum bicolor	19*
Oryza sativa	22
Selaginella moellendorffi	3*
Physcomitrella patens	10
Chlamydomonas reinhardtii	5
Chlorella vulgaris	3*
Cyanidioschyzon merolae	2
Homo sapiens	11
Mus musculus	10
Gallus gallus	10
Xenopus tropicalis	11
Danio rerio	14
Branchiostoma floridae	11
Ciona intestinalis	8
Drosophila melanogaster	7
Anopeheles gambiae	6
Aedes aegypti	6
Caenorhabditis elegans	7
Schizosaccharomyces pombe	2
Dictyostelium discoidum	2
Plasmodium falciparum	3
Phytophthora sojae	3

Table 3: The number of SR family members in various species (Richardson et al., 2011). The numbers marked with asterisk might be lower than the actual number of SR proteins in that particular species, because the sequences without methionine at the beginning were omitted.

The well established role of SR proteins lies in assistance to general splicing machinery in recognition of correct splice sites in pre-mRNA. RS domain is crucial for recruitment of components of U1 and U2 snRNPs. It is heavily phosphorylated and serves as a surface for protein-protein interaction with other splicing factors also containing RS domains like U1-70K and U2AF35, proteins associated with U1 and U2 snRNPs. In recent years many other functions during the mRNA life cycle were attributed to SR proteins. They promote export of mRNA from the nucleus by interaction with TAP export receptor, they stimulate transcription elongation, NMD (nonsense-mediated decay) pathway, translation and have mRNA unrelated functions like maintaining genome stability, enhancing protein sumoylation and some of them are proto-oncogenes (Long and Caceres, 2009; Twyffels et al., 2011; Zhong et al., 2009). SR proteins are essential and nonredundant at least in some tissues and developmental stages. Knockout mice for several SR proteins had all an early embryonic phenotype. SRSF1 is also essential for *Caenorhabditis elegans* development (Twyffels et al., 2011).

SR proteins are predominantly nuclear with a tendency to accumulate in nuclear speckles, which is a storage compartment of splicing and transcription factors. SR proteins accumulate in speckles, while their RS domain is dephosphorylated and when it is rephosphorylated, they enter nucleoplasm and bind pre-mRNA. Most of SR proteins are also able to shuttle between nucleus and cytoplasm, and the RS domain serves as their nuclear localisation

signal (Twyffels et al., 2011). Reversible phosphorylation of the RS domain is key for SR protein versatility in their cellular functions.

Just like the general splicing machinery, SR proteins are widespread among eukaryotes and their number varies among organisms. The eukaryotic taxa that completely lack SR proteins, lack them secondarily, because the common ancestor of eukaryotes presumably had SR proteins. The number of SR proteins is related to splicing complexity, alternative splicing and conservation of splice site and branch point sequences in each particular group (**Table 3**). The presence of SR proteins relieves the selective pressure on maintaining the sequence of splice sites close to consensus, which leads to splice site divergence and more alternative splicing (Busch and Hertel, 2012). It can be demonstrated on yeasts. The budding yeast *Saccharomyces cerevisiae* that has introns in about 2.5 % of genes and lacks alternative splicing, has no SR proteins (it possess some SR-like proteins), whereas the fission yeast *Schizosaccharomyces pombe* that has forty times more introns, has two SR proteins. Higher plants have the most SR proteins and even some plant-specific subfamilies. They acquired the extra members of the SR family presumably due to frequent genome duplications. It is a question though, if all their SR genes are essential, because it was described that one of the paralogs arising from the whole genome duplication is always expressed less and therefore probably redundant (Reddy and Shad Ali, 2011). The research on plant SR proteins is lagging behind the animal field so there is, among other things, no genome-wide information on their binding to RNA or acetylation, which was my prime interest, so we cannot draw here a comparison to animals.

Binding of SR proteins to RNA

The RNA recognition motive of SR proteins is responsible for their sequence-specific binding to RNA. They recognize only about four nucleotides long purine-rich sequences so the RNA sequence cannot be the sole determinant of SR protein binding. Very important is protein-protein interaction via RS domain and its phosphorylation state (Shepard and Hertel, 2009). The SR binding site is usually degenerated, so the RRM must be able to bind various RNA sequences. This is achieved either by versatility of RRM that can accomodate various nucleotides or by the fact that different RNA sequences are bound in distinct conformations (Anko and Neugebauer, 2012). Mutation in the SR binding site is a common cause of disease (Cartegni et al., 2002) as demonstrated, e.g., for SRSF1 (Sanford et al., 2009).

Thanks to recent development of high-throughput methods, binding of SRSF1 (Sanford et al., 2009), SRSF3 and SRSF4 to RNA (Anko et al., 2010; Anko et al., 2012) was studied on genome-wide scale *in vivo* by RIP-chip (RNA immunoprecipitation followed by hybridization to microarray) or CLIP-seq (photocrosslinking and IP followed by next-generation sequencing). CLIP is a unique method, that provides a snapshot of true situation *in vivo*. Regarding the consensus binding sequence, it confirmed older *in vitro* data so on this occasion the *in vitro* approach (SELEX) was valid and led to meaningful results.

SRSF1 bound predominantly exonic sequences. Its binding sites were enriched in constitutive exons next to the alternative ones, which suggests a mechanism of alternative exon skipping modulated by SRSF1 due to activation of downstream splice site in constitutive exon. A significant group of mRNAs among the SRSF1 targets were RNA binding proteins participating in all steps of gene expression (Sanford et al., 2009).

For SRSF3 and SRSF4 the CLIP study challenged the accepted model, that SR proteins predominantly bind to splicing regulatory sequences in exons (so called exonic splicing enhancers) and promote recognition of the adjacent splice site, because majority of SRSF3 and SRSF4 were bound to intronic sequences and it is even more striking when you consider that in cells there is much more mRNA than pre-mRNA that SR proteins can bind to. So for SRSF3 and SRSF4 the splicing regulatory elements in introns are at least as important as those in exons.

Each SR protein has several thousand targets in the cell and the targets of individual SR proteins overlap only slightly. This supports the view, that each SR protein regulates subset of splicing events in the cell and therefore the differences in expression level of splicing regulators in each tissue might be the key factor that causes tissue-specific alternative splicing. SRSF3 and SRSF4 bound also many non-coding RNAs. Most significant was probably MALAT1, RNA enriched in nuclear speckles, that was previously shown to bind SR proteins (Tripathi et al., 2010), then 7SK RNA, component of P-TEFb and snoRNAs (small nucleolar RNAs) particularly scaRNAs (small Cajal body-specific RNAs), that are processed from introns so SR proteins may play some role in snoRNA maturation. SRSF3 and SRSF4 bound also mRNAs from intronless genes like polyadenylated histone mRNAs in their 3′ UTR. So for histone mRNAs, SR proteins are important in further steps of their expression, probably for export from the nucleus and translation. SRSF3 was revealed as a regulator of other SR proteins. It was known previously, that each SR protein regulates its quantity in the cell by

promoting the inclusion of alternative exon containing premature termination codon into its mRNA. Such exon then triggers nosense-mediated decay of that RNA (Lareau et al., 2007; Ni et al., 2007). SRSF3, apart from itself, regulated SRSF2, 5 and 7. It bound also other RNA binding proteins, splicing factors and proteins from the spliceosome, which makes it an important modulator of RNA metabolism (Anko et al., 2012).

So far the structure of a complete SR protein bound to RNA has not been resolved, because RS domain is intrinsically disordered, heterogeneously phosphorylated and insoluble and therefore cannot be visualized by crystallography or NMR. Yet there are structures of RRM domain of SR proteins SRSF1, 2, 3 and 7 in complex with short RNA oligonucleotides (Clery et al., 2011; Hargous et al., 2006; Phelan et al., 2012; Tintaru et al., 2007) that reveal which residues of RRM are involved in RNA binding and also uncover structural basis of recognition of RNA sequence by RRM. In general RNA is bound by an interaction surface formed by several β-sheets that are supported from the opposite side by two α-helices. Two α-helices and several β-sheets oriented as described are a canonical fold of an RRM, but this fold can be contacted by RNA in different ways.

The proteins used for structural studies did not carry acetylations or other eukaryotic posttranslational modifications, because they were purified from bacteria. The lysines that were reported to be acetylated *in vivo* were not directly contacting RNA, but amino acids in close proximity were essential for RNA binding. Interestingly the amino acids directly binding RNA are not more conserved than the surrounding residues and in general there are more modes of RNA binding among RRM domains. They differ in involvement of certain loops between β-sheets of RRM and contribution of amino acids on the N- and C- termini of the RRM to RNA binding. The structure of SRSF1 RRM2 is the closest

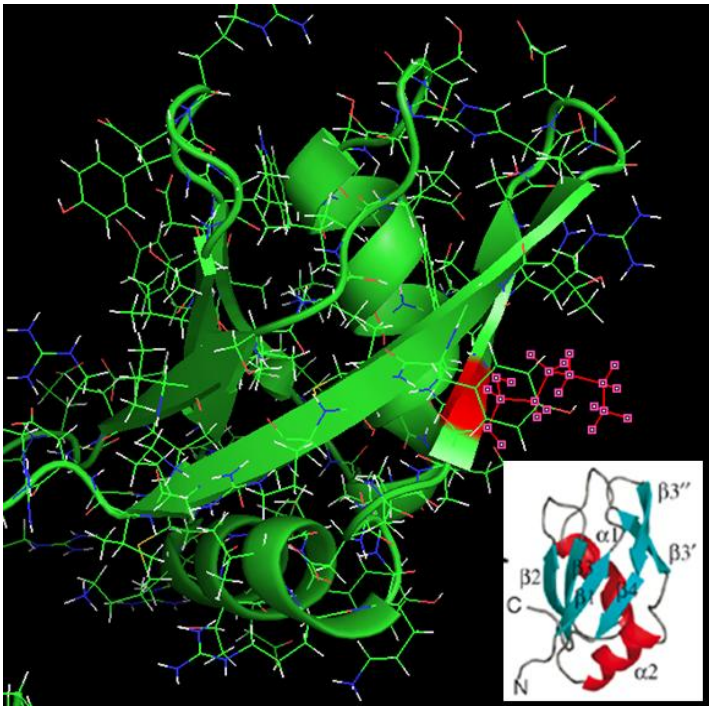

Fig. 1: The structure of SRSF1 RRM2 with the acetylated lysine residue highlighted in red (Tintaru et al., 2007). PDB entry 2O3D, visualization done in PyMOL.

approximation of so far unknown structure of SRSF5 RRM2, which is the subject of my study. In SRSF1 the acetylated lysine is located in β-sheet β3′ (Tintaru et al., 2007), where it can potentially regulate RNA binding (**Fig. 1**).

.

Acetylation of SR proteins and its function

Majority of the 7 classic SR proteins in humans are acetylated in the RRM domain with the exception of SRSF4 that is not acetylated and SRSF6 that is acetylated outside the RRM domain. SRSF7 is acetylated in RS domain as well (**Fig. 2**). Moreover, acetylation sites are conserved in related SR proteins. The identical lysine residue is acetylated in SRSF3 (K23) and SRSF7 (K24), similarly SRSF1 (K38) and SRSF2 (K36), and SRSF5 (K167) and SRSF1 (K179). In the newly included SRSF8 – SRSF12 no acetylation has been detected so far, but it might due to the fact that they are generally less studied.

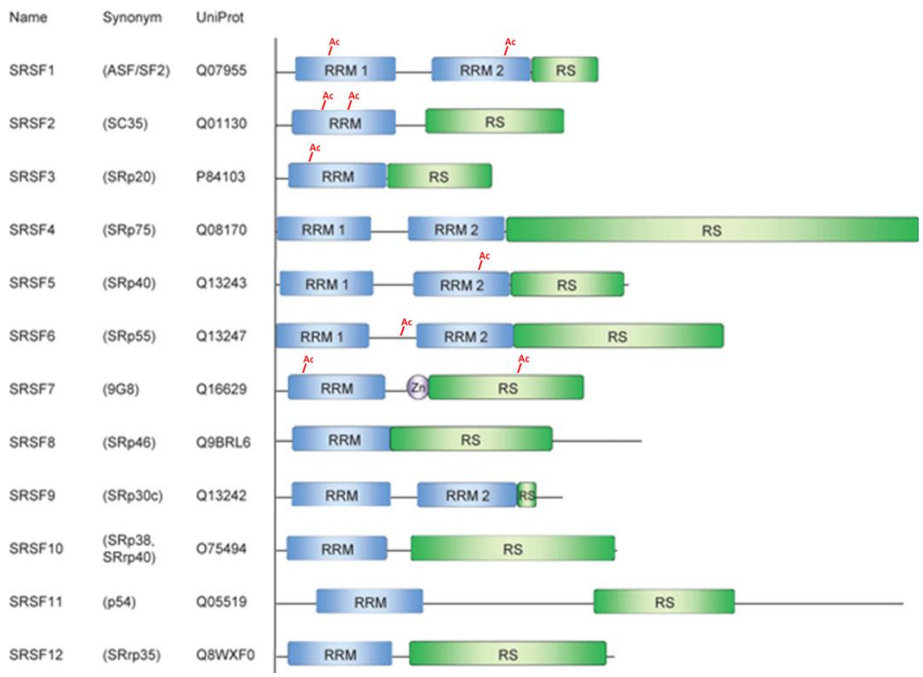

Fig. 2: The family of human SR proteins with the acetylation sites marked. Edited from (Twyffels et al., 2011).

SRSF2 is the only member of SR protein family, whose acetylation was investigated (Edmond et al., 2011b). Acetylation by Tip60 had a destabilizing effect on SRSF2 and promoted its proteasomal degradation in human lung carcinoma cell line. Mutation of the acetylated lysine to arginine stabilized the protein. In agreement with the previous observation, HDAC6 mediated deacetylation stabilized SRSF2 as well. Often, acetylation has stabilizing effect on proteins by inhibiting ubiquitination of the same lysine residue, but it is well established that acetylation can have destabilizing effect on proteins too. It is not known, whether SRSF2 is degraded by ubiquitin-dependent or independent pathway. Tip60 has also an indirect effect on SRSF2 phosphorylation. Overexpression of Tip60 caused relocation of SR protein kinases to the cytoplasm and subsequent SRSF2

hypophosphorylation. Phosphorylation is required for SR protein function so Tip60 is a negative regulator of SRSF2 on more levels. It causes destabilization as well as dephosphorylation of SRSF2 (Edmond et al., 2011b). Finally they showed that after treating cells with cisplatin, SRSF2 accumulates in nonacetylated and phosphorylated form. However, the role of SRSF2 acetylation in interaction with RNA was not analyzed and currently there are no data avialble that would probe role of SR protein acetylation in their RNA binding. Disturbingly, Edmond and colleagues reported acetylation of different lysine residue than was previously described (Choudhary et al., 2009).

A follow-up study from the same lab assessing the influence of HDAC inhibitor sodium butyrate (NaBu) and SRSF2 on senescence (Edmond et al., 2011a) came to different conclusions. In human lung carcinoma cell line treated with NaBu, SRSF2 accumulated in acetylated and hypophosphorylated form. It is unclear why the acetylated SRSF2 is stabilized in this case. Other HDAC inihibitors trichostatin A (TSA) and nicotinamide (NAM) had no effect on SRSF2 acetylation even though at least TSA should inihibit HDAC6 previously shown to deacetylate SRSF2. This confusing result can be explained by the finding, that in this case NaBu did not function as an HDAC inhibitor, but it rather upregulated the Tip60 acetyltransferase. TSA and NAM did not change Tip60 expression (Edmond et al., 2011a). Interestingly, our recent data showed that NaBu did not change acetylation of another SR protein SRSF5 (Hnilicova et al., 2011).

In this work I developed and tested a hypothesis, that lysine 167 of SRSF5 and its acetylation is an important regulatory modification, that can influence binding of SRSF5 to RNA, splicing of SRSF5 targets and also the SRSF5 subcellular localisation or phosphorylation of its RS domain. To test it, I mutated the putative acetylation site of SRSF5 to arginine and glutamine. The arginine mutation mimics nonacetylated lysine, whereas the glutamine mutation should imitate acetylated lysine. To distinguish mutants from endogenous proteins I tagged the exogenous proteins with EGFP. I tested various properties of those mutants in comparison to the wild type protein. Their binding to RNA *in vitro* by fluorescence anisotropy and in HeLa cell lines by RIP. Their localisation in the cell by immunofluorescence microscopy and their dynamics by fluorescence recovery after photobleaching (FRAP), the acetylation and phosphorylation state of the RS domain by western blot and a direct effect of mutants on splicing by splicing reporter containing SRSF5 binding site.

Materials and methods

Cell culture and drugs

HeLa cells were cultured in DMEM high glucose medium (Sigma) supplemented with 10% fetal calf serum and penicillin/streptomycin. Cells were grown in an incubator at 37 °C in 5% CO2 atmosphere. The SRSF5-EGFP BAC HeLa cell line was kindly provided by Karla Neugebauer, MPI-CBG, Dresden.

The inhibitor of transcription elongation, 5,6-Dichloro-1-β-D-ribofuranosylbenzimidazole (DRB) (Sigma) was used at 50 µM final concentration in the medium. Cells were incubated with DRB for 4 hours. The inhibitor of HDACs, sodium butyrate (NaBu) (Sigma) was used at 5mM final concentration in the medium and cells were treated for 6 hours.

Protein tagging and mutagenesis

The SRSF5 gene was amplified from mouse cDNA using primers F 5′–CCC AAG CTT ATG AGT GGC TGT CGA GTG–3′ and R 5′–CCG GAA TTC CAT TGC CAC TGT CAA CTG A–3′ and cloned into EGFP-N1 vector (Clontech) using EcoRI and HindIII restriction sites. The accuracy of the construct was checked by restriction digest and by sequencing.

The point mutation was created by PCR mutagenesis. It was performed in 50 µl PCR reactions using Phusion polymerase (Thermo Scientific), 20 ng plasmid DNA template and 100 nM primers. Two sets of primers were used with the one nucleotide mismatch in the middle. For the mutation of Lys to Arg F 5′–GCT ATT GAG AAA CTT TCT GGA AGG GAA ATT AAC GGG–3′ R 5′–AGA AAA ATC GAT TTT TCT CCC GTT AAT TTC CTT TCC AGA AAG TTT CTC AAT AGC–3′ and for the mutation of Lys to Gln F 5′–TGC TAT TGA GAA ACT TTC TGG ACA GGA AAT TAA CGG GAG AAA AAT–3′ R 5′–ATT TTT CTC CCG TTA ATT TCC TTT CCA GAA AGT TTC TCA ATA GCA–3′. PCR reaction was performed according to the Phusion polymerase manufacturer`s recommendations with primer annnealing temperature 65 °C for 18 cycles. After cleanup of the reaction by QIAquick PCR purification kit (Qiagen) the template DNA was cleaved by methylation-sensitive restriction enzyme DpnI (NEB) overnight at 37 °C. The resulting mixture was used for heat shock transformation of DH5α cells. Accuracy of the construct was checked by sequencing and the SRSF5 gene sequences were recloned into fresh EGFP-N1 vectors that did not undergo the mutagenesis.

Transfection of siRNA, plasmid DNA and creating stable cell lines

Transfection of siRNA and DNA was done as previously described (Hnilicova et al., 2011). The siRNA against SRSF5 (Silencer® Select Ambion) sequence: sense 5′–GAA UUA GUU UA AUG CCU UAtt–3′ and antisense 5′–UAA GGC AUU AAA CUA AUU Ctg–3′. As a negative control I used Negative Control#1 siRNA (Ambion).

For production of stable cell lines expressing WT or mutant SRSF5-EGFP, cells were plated on a 10 cm Petri dish and after 24 hours transfected with 5 μg of appropriate plasmid. After another 24 hours G418 antibiotic was added to the culture medium at 1 mg/ml to kill the cells not expressing the construct. Culture medium with fresh antibiotic was exchanged every 48 hours and after two weeks only colonies of stably transformed cells remained on the dish. Cells were then sorted on FACS and only the ones exhibiting higher fluorescence were used in further experiments.

Antibodies and western blotting

The Acetylated-Lysine Antibody (Cell Signaling technology, rabbit polyclonal) was diluted 1:250 in 5% milk dissolved in PBS supplemented with 0.05% Tween and incubated with the membranes O/N at room temperature. The supernatant of m104 hybridoma (mouse monoclonal antibody recognizing a phosphorylated epitope on SR proteins, kindly provided by K. Neugebauer, MPI-CBG Dresden) was diluted 1:10 in 5% BSA dissolved in TBS supplemented with 0.1% Tween and 25mM NaF and incubated with the membranes for 30 minutes at room temperature. The GFP antibody B-2 (Santa Cruz Biotechnology, mouse monoclonal) was diluted 1:1000 in 5% milk dissolved in PBS supplemented with 0.05% Tween and incubated with the membranes for 1 hour at room temperature. Anti-GAPDH G9545 antibody (Sigma, rabbit polyclonal) was diluted 1:1000 in 5% milk dissolved in PBS supplemented with 0,05% Tween and incubated with the membranes 1 hour at room temperature. Appropriate secondary antibodies conjugated with HRP (Jackson ImmunoResearch) were used at 1:5000 dilution and incubated with the membranes for 1 hour at room temperature. The western blots were developed with the ECL substrates SuperSignal West Femto and Pico Substrates (Thermo Scientific). The results were quantified in ImageJ.

Immunofluorescence

Cells were grown on cover slips to 30–50% confluency. Between each of the following steps they were washed 3 times with PBS. Cells were fixed in 4% paraformaldehyde/PIPES for 10 minutes, then permeabilized with 0.5% TritonX-100 (Sigma) in PBS for 5 minutes and blocked with 5% normal goat serum (Jackson ImmunoResearch) for 15 minutes. Next the supernatant of mouse hybridoma recognizing SC-35 (kindly provided by K. Neugebauer, MPI-CBG Dresden) was diluted 1:10 in 5% normal goat serum and the cover slips were incubated on a drop of antibody for 1 hour in a wet chamber. Then they were incubated with a secondary antibody diluted 1:200 conjugated with DyLight549 (Jackson ImmunoResearch) for 1 hour in a wet chamber. Finally they were washed in dH_2O, dried and mounted into Fluoromount (SouthernBiotech) with DAPI. Images were acquired as described previously (Huranova et al., 2009) and the deconvolution was performed in Huygens software package (Scientific Volume Imaging).

Splicing reporter assay, RNA isolation and RT-PCR

RNA was isolated from 80–100% confluent cells grown on one well of 12-well plate for 3 days. One day after plating, cells were transfected with siRNA against SRSF5 or with negative control#1 siRNA (Ambion). After one more day cells were transfected with 500 ng of splicing reporter. Cells were harvested 24 hours after transfection of splicing reporter and RNA was isolated by TRIzol reagent (Ambion) according to manufacturer`s protocol. 1 μg of RNA was used in RT reaction with SuperScript III (Invitrogen) according to manufacturer`s protocol. cDNA from the splicing reporter was amplified by Taq polymerase (Fermentas) with primers F 5′–TGG AGT ACA ATG TCA GTG TTT–3′ and R 5′–CTG GAC CAA TGT TGG TGA ATC–3′. The amount of SRSF5 mRNA was assessed by qPCR where different forward primers distinguished between endogenous and exogenous mRNA and the reverse primer was the same for both variants. The specificity of primers was confirmed with mouse and human cDNA. Endogenous human F 5′–GGA TGC AGA TGA TGC TGT GT–3′ mouse F 5′–GGA TGC AGA TGA TGC TGT TT–3′ common R 5′–TCA TTT CGA GGT CTG GGA CT–3′.

Protein expression and purification

Two RRM domains (first 180 amino acids) of SRSF5 were cloned into pET-42b vector using NcoI and NotI restriction enzymes to create a fusion protein tagged with GST on N-terminus. Following primers were used F 5′–CAT GCC ATG GAT GAA AAC CTG TACT TCCA GGG CAT GAG TGG CTG TCG AGT G–3′ R 5′–ATA GTT TAG CGG CCG CTT ATT TGC TGC CTT CAA TTA A–3′. BL21 bacterial cells (Invitrogen) were transformed with the plasmid and grown in LB medium with 25 µg/ml kanamycin. Protein expression was induced at OD 0.5–1 with 250 µM IPTG. After shaking for 3 hours at 37 °C cells were spun, rinsed with dH2O and stored in -80 °C. Pelets were resuspended in GST loading buffer (PBS supplemented with 5 mM dithiothreitol (Sigma), 0.1% Triton X-100 (v/v) (Sigma) and protease inhibitor cocktail (Calbiochem)). Cells were sonicated and the lysate was cleared by centrifugation at 20,000 g for 30 minutes at 4 °C. Lysate was incubated rotating with glutathione agarose beads (Sigma) for 3 hours at room temperature. Beads were transferred onto a column and washed 8 times with 15 ml of GST loading buffer and 2 times with 15 ml of Elution buffer (Tris pH 8.0, 50 mM NaCl). GST fusion protein was eluted with Elution buffer supplemented with 10 mM glutathione (Sigma).

Protein concentration was determined by standard Bradford assay using 600 µl Bradford reagent (Sigma). BSA standards from 200 to 1000 µg/ml were used for measurement of the calibration curve. Protein purity was assessed by SDS-PAGE on 10% gels and PageBlue Protein Staining Solution (Fermentas) was used for gel staining. To exchange buffers and concentrate the proteins, samples were repeatedly spun at 15,000 g in 500 µl Vivaspin columns (Sartorius) with a polyethersulfone membrane and 10 kDa molecular weight cutoff.

Thermofluor measurement

The thermofluor measurement is a fluorescence-based thermal stability assay (Pantoliano et al., 2001). Thermal stability of a protein is measured by incubating the protein with hydrophobic fluorescent probe that is quenched in aqueous solution while steadily increasing the temperature. As the protein unfolds, the hydrophobic residues from its core interact with the fluorophore, which is no longer quenched and fluorescence of the sample increases. The protein transforms from folded to unfolded state rapidly around a particular temperature and the midpoint of this transition is defined as the melting temperature, Tm (Ericsson et al., 2006).

The protein Tm was assessed in pannel of buffers (Ericsson et al., 2006) in 96 well plate format. The purified, concentrated protein was diluted into tested buffers to a 50 µg/ml final concentration and SYPRO Orange Protein Gel Stain (Invitrogen) was added at final concentration 8×. The total volume of each sample was 25 µl. 24 buffering agents of varying pH with addition of either 200 mM or 400 mM NaCl were tested. Fluorescence was measured in Light Cycler 480 (Roche) with temperature increasing from 20 to 95 °C at a ramp rate of 0.01 °C/s.

Fluorescence anisotropy

The equilibrium binding of SRSF5 protein to RNA was analyzed by fluorescence anisotropy. Measurements were conducted on a FluoroMax-4 spectrofluorometer (Horiba Jobin-Yvon) instrument equipped with a thermostatted cell holder with a Neslab RTE7 water bath (Thermo Scientific). The system was operated using FluorEssence V3.5 software (Horiba Jobin-Yvon). The 5′-fluorescein labeled RNA probe 5′–UGA CUC CAC UGC–3′ (Sigma) was designed as two consecutive SRSF5 consensus binding sites based on literature (Caputi et al., 2004; Goren et al., 2006; Liu et al., 1998; Tacke et al., 1997). The RNA probe was excited at 488 nm and its emission was recorded at 520 nm. The widths of monochromatic slits were 9 nm for both excitation and emission and integration time was set to 3 s. The measurements were done at 10°C in 100 mM bicine buffer (pH 7.6) containing 300 mM NaCl, 0.5 mM DTT and 10% glycerol. 10 nM labeled RNA (1.4 ml) was titrated with increasing amounts of purified protein and the reaction was carried out in a stirred 1.5 ml quartz cuvette. Each data point is an average of three measurements. The experimental data were analyzed in SigmaPlot 11 software (Systat Software) and isotherms were fit to a single-site binding model according to Heyduk and Lee using nonlinear least squares regression. The data were normalized for visualization purposes.

FRAP

For FRAP measurements, cells were plated on glass bottom dishes (InVitro Scientific) and measured 24 hours later at 30–50% confluency. Prior to FRAP measurements, the medium was exchanged for colorless DMEM high glucose medium (GIBCO). Cells were imaged on Leica TCS SP5 confocal microscope, which is equipped with a chamber for live cell imaging. Objective HCX Plan-Apochromat 63× NA 1.40-0.6 oil, Lbd Blue CS was used. Images were

acquired with open pinhole (250 µm) in 12-bit resolution in 512 × 512 format at a 1000 Hz scan speed (imaged every 0.277 s). The 488 nm laser line of argon laser (100 mW) at 6% laser power was used for imaging and 458 nm, 476 nm, 488 nm, and 514 nm lines were used for bleaching at 100% laser power. In each measurement, 12 prebleach images were taken, followed by 3 bleach pulses and 400 postbleach images. In all the experiments half of the nucleus was bleached. This method was described previously (Phair et al., 2004a; Phair et al., 2004b) and is suitable for fast moving, transiently interacting nuclear proteins. Around 10 cells were measured in each experiment and the measured fluorescence was normalized to the whole nucleus fluorescence and background was subtracted. The FRAP curves were fitted with double exponential function using SigmaPlot and the half-times of fluorescence recovery were extracted.

Immunoprecipitation and RNA immunoprecipitation (RIP)

Immunoprecipitation was performed as described previously (Huranova et al., 2009). RNA immunoprecipitation was performed from 90–100% confluent PD15 of cells grown for 48 hours. The cell lines expressing SRSF5-GFP were used and normal HeLa cells not expressing GFP served as a negative control. All the steps were performed on ice or at 4 °C. Cells were washed three times with cold PBS, scraped into 1 ml of cold PBS and centrifuged at 1000 g for 2 minutes. Cell pellets were resuspended into RIP lysis buffer (20 mM Tris-HCl pH 7.5, 200 mM KCl, 2 mM $MgCl_2$, 1% (v/v) NP-40) to the mass of 1.5 g and 7.5 µl protease inhibitor cocktail (Calbiochem) and 3 µl RNasin (Promega) were added. The solution was sonicated in 0.5 s pulses and centrifuged at 20,000 g for 10 minutes. The supernatants were transferred into new tubes and 30 µl of Sepharose G beads (GE Healthcare) prewashed in RIP lysis buffer were added for precleaning and the solutions were incubated rotating at 4 °C for 1 hour. During the precleaning their protein concentration was measured and all the samples were diluted afterwards to identical protein concentration with RIP lysis buffer. Ten percent of the samples were put in the freezer as input.

7.5 µg of anti-GFP antibody (obtained from David Drechsel, MPI-CBG, Dresden, Germany) was added to all samples and they were incubated rotating at 4 °C for 2 hours. Samples were centrifuged at 10,000 g for 5 minutes to get rid of the potential protein precipitate and supernatants were mixed with 40 µl of Sepharose G beads preblocked with 1000 µg/ml BSA (NEB) and 100 µg/ml yeast RNA (Ambion) and incubated rotating at 4 °C for 1 hour. The

beads were washed 6 times with RIP lysis buffer and one quarter of them was saved for protein analysis by western blot and the rest was mixed with 300 µl of NET-2 buffer (50 mM TRIS–Cl pH 7.5, 150 mM NaCl, 0.05% Nonidet P-40) supplemented with 1% SDS and 300 µl of acid phenol:chloroform (Sigma) and incubated shaking at 37 °C for 1 hour. From this step of RNA isolation the input samples were processed as well. The mixture was centrifuged at 10,000 g for 10 minutes at room temperature and the aqueous phase was transferred into a new tube and RNA was precipitated with 1 ml of 96% ethanol, 30 µl of 3 M sodium acetate and 20 µg of glycogen (Ambion) overnight in -80 °C.

The precipitated RNA was resuspended in 7.5 µl of nuclease-free water and the trace DNA was eliminated by DNAfree kit (Ambion) according to manufacturer`s instructions. 4 µl of RNA was used in reverse transcription with gene-specific reverse primers later used in qPCR and 4 µl of RNA was stored as a -RT control for qPCR. The RT reactions were diluted 5 times with PCR H_2O and served as a template in qPCR reactions. Light Cycler® 480 SYBR Green I Master mix (Roche) was used according to manufacturer`s instructions. The qPCR was run in Light Cycler® 480 (Roche) with annealing temperature 61 °C for 45 cycles. All primers were tested for efficiency and only those with efficiency between 90 and 110% were used. The specificity of primers was always checked by melting curve analysis. The following primer pairs were used at 1 µM concentration. SRSF5 F 5′–AGA GTC AGC TGG CAG CCT GTC TGT G–3′ SRSF5 R 5′–ATC ACT GTA GGA GCT GAC TGG CAA A–3′ from (Lareau et al., 2007) and for checking that the RNA concentration in inputs is the same, I used common GAPDH primers F 5′–GAA GGT GAA GGT CGG AGT–3′ R 5′–GAA GAT GGT GAT GGG ATT TC–3′. The RIPs were quantified as percent of input normalized to GFP signal from the RIP protein fraction. This was necessary, because each cell line had slightly different level of expression of SRSF5-EGFP and different amounts of SRSF5-EGFP were pulled down.

Results

First step in studying the acetylation of SRSF5 was confirming the mass spectrometry data about the acetylation (Choudhary et al., 2009) on western blot. I had available a HeLa cell line with genome-integrated bacterial artificial chromosome containing full-length mouse SRSF5 gene tagged with EGFP with its own promoter and other regulatory sequences (Poser et al., 2008). I called it WTBAC. The cells express the protein at physiological level. It was shown that such proteins behave in all aspects as the endogenous ones (Anko et al., 2010). I performed immunoprecipitation from the cells via the EGFP tag and I confirmed the acetylation on western blot (**Fig. 3**). I also tested if SRSF5 acetylation is responsive to sodium butyrate (NaBu), a broad spectrum inhibitor of HDACs, and found out that SRSF5 acetylation does not change after NaBu treatment (**Fig. 3**).

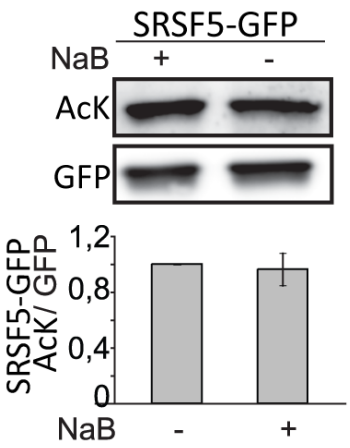

Fig. 3: Acetylation of SRSF5-GFP in cells treated with sodium butyrate (NaBu) and in nontreated cells. AcK, anti-acetylated lysine antibody. Acetylation was quantified by denzitometric analysis in ImageJ. The western blots with AcK antibody were normalized to the total amount of immunoprecipitated protein inferred from GFP signal; average of four biological experiments; error bars represent standard deviation. The acetylation in nontreated cells was set as 1. Adapted from (Hnilicova et al., 2011).

To study the protein further, I cloned the SRSF5 cDNA from mouse 3T3 cells into EGFP vector to tag the protein with EGFP on C-terminus and mainly for subsequent mutagenesis. Mouse SRSF5 has 270 amino acids (31 kDa) and is on protein level almost identical to the human protein apart from 6 conservative substitutions or indels in the variable RS domain. I chose the strategy of introducing the mouse protein into human cells for future simplier knockdown of endogenous protein and because I wanted to use the WTBAC as a positive control in certain experiments, which is mentioned later.

SRSF5 has single acetylation site at lysine 167 in the RRM2 domain. To study how changes of amino acid side chain in this position alter protein function, I created two different point mutations of K167 by PCR site-directed mutagenesis (**Fig. 4a**). The mutation to arginine preserves positive charge at that position, but blocks acetylation. The mutation to glutamine, on the other hand, mimics acetylation. I named them R and Q respectively and compared various properties of the mutants with wild type SRSF5.

To control for the influence of EGFP tag on SRSF5 function, I included a positive control The WTBAC protein, which is several kDa heavier because it contains a longer linker between

SRSF5 and EGFP than the other three proteins, hence the difference in size on western blot (**Fig. 4b**).

SRSF5 localisation

I created HeLa cell lines stably expressing WT, R and Q variants of SRSF5 at similar levels close to physiological and I observed localization of the proteins in cells (**Fig. 4a**). All the

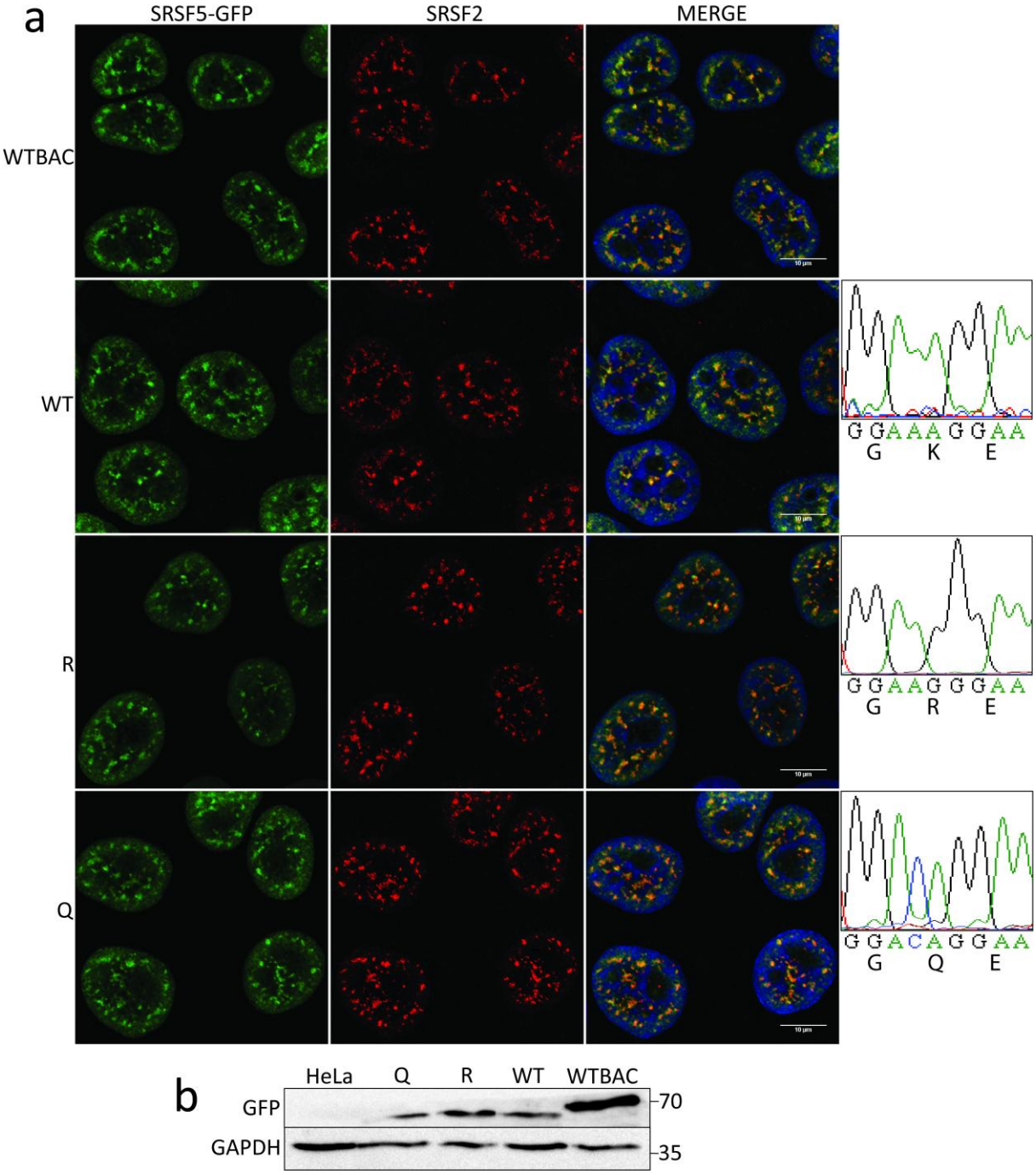

Fig. 4: (a) Subcellular localisation of WT and mutant variants of SRSF5-GFP and their colocalization with a speckle marker SRSF2 and DNA counterstained with DAPI (blue), scale bar 10 µm. On the right, the point mutations confirmed by DNA sequencing. **(b)** The expression level of SRSF5-GFP in the lysates from the cell lines probed with GFP antibody; GAPDH served as a loading control and HeLa cells not expressing GFP as a negative control.

proteins had, as expected, nuclear localization and they accumulated in certain nuclear regions that were confirmed to be nuclear speckles by staining against a nuclear speckle marker SRSF2. The positive control WTBAC was identically distributed. The natural localization of WT, R and Q indicated that they were functional in splicing and other processes they participate in. The localization of SRSF5 was transcription-dependent. After inhibition of transcription by DRB, SRSF5 accumulated in enlarged speckles (not shown), which is a general property of splicing factors.

The lack of difference in subcellular localization among WT, R and Q was surprising (**Fig. 4a**). It indicated that even the proteins with the mutation are functional and if there is some difference among them, it does not manifest itself on protein localization.

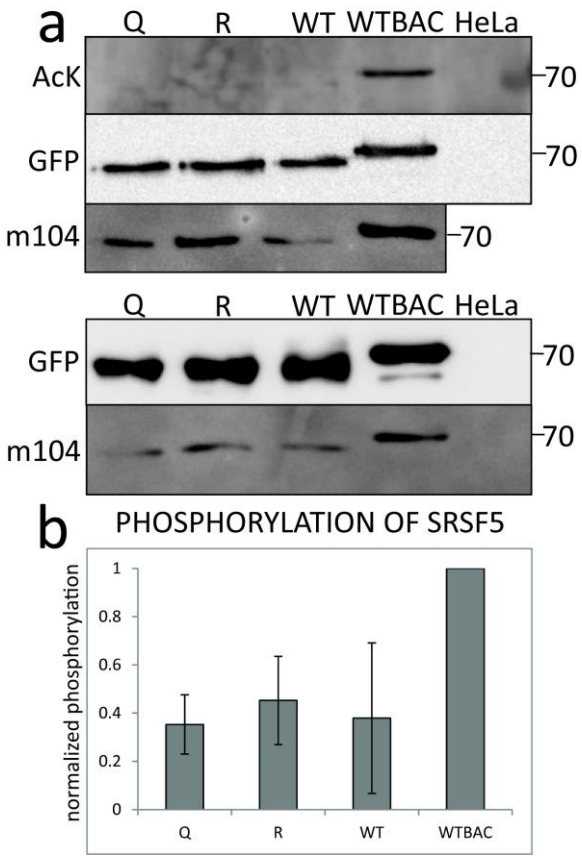

Fig. 5: (a) Acetylation and phosphorylation of immunoprecipitated SRSF5-EGFP; HeLa denotes a negative control immunoprecipitation from standard HeLa cells without GFP expression; AcK, anti-acetylated lysine antibody; m104, anti-phosphorylated SR protein antibody **(b)** Denzitometric analysis in ImageJ of western blots with m104 antibody normalized to the total amount of immunoprecipitated protein inferred from GFP signal; average of two biological experiments; error bars represent standard deviation. The WTBAC phosphorylation was set to 1.

Acetylation and phosphorylation of SRSF5

Next, I examined the posttranslational modifications of the four studied SRSF5-EGFP proteins. I immunoprecipitated them with anti-GFP antibody and probed the modifications on western blot with anti-acetylated lysine antibody (AcK) to confirm that the mutants are not acetylated and anti-phosphorylated SR antibody (m104) to see the level of phosphorylation in their RS domain (**Fig. 5a**). I was able to reliably detect the acetylation of WTBAC protein unlike the WT protein. In case of WT and also R protein, there was a very faint signal that is hard to interpret. It could come from acetylation site other than K167, however, if that is the case, similar signal should be detected in the Q mutant as well, but there was none. Alternatively, it is unspecific binding of the AcK antibody. The reason why WT protein

was not acetylated like the WTBAC remains an open question because the proteins differ only in the linker sequence between SRSF5 and EGFP.

The WT, R and Q proteins were similarly phosphorylated; there was not a statistically significant difference among them. The WTBAC was phosphorylated two to three times more (**Fig. 5b**). The presence of phosphorylation again proved that the proteins are functional. The higher phosphorylation of WTBAC did not manifest in its different localization, even though it is well established that SR protein phosphorylation is one of determinants of their subcellular localization. Probably the difference was not big enough to be perceptible. Also we must not forget that the tagged protein is competing with the endogenous protein for binding sites, which can influence its localization as well.

Dynamics of SRSF5 proteins *in vivo*

I measured SRSF5 dynamics in the cell by FRAP. In the experiment, half of the nucleus was bleached (**Fig. 6c**), because I wanted to cover the proteins in nucleoplasm and speckles at the same time and also because the fluorescence recovery was too fast to measure correct recovery curve when I bleached a small circular area. The FRAP curves I gained were very similar for all four SRSF5 proteins (**Fig. 6a**). The fluorescence recovery was very fast and there was no or very small immobile fraction. Because each cell line bleached with slightly different efficiency, I used the half time of fluorescence recovery to compare the FRAP curves with each other. Half times were not significantly different among the proteins.

This method allows to study the interaction of SRSF5 proteins with RNA *in vivo*, because the dynamics of SR proteins in cells are influenced by two factors, their diffusion and binding to RNA or proteins. The diffusion should be the same for the four SRSF5 proteins I studied. Molecular weight of WT, R and Q is almost the same, because they differ only in one amino acid substitution and the WTBAC is only a few kDa heavier, which is negligible, if you take into account that you measure mixed population of proteins that are differentially phosphorylated and acetylated. The mutation in RRM2 in R and Q proteins could alter their interaction with RNA but should not change interaction with proteins, via the RS domain. So the observed difference in fluorescence recovery should be caused by altered interaction with RNA. The conclusion is that I was not able to detect differential interaction of SRSF5 proteins with RNA *in vivo* by this method.

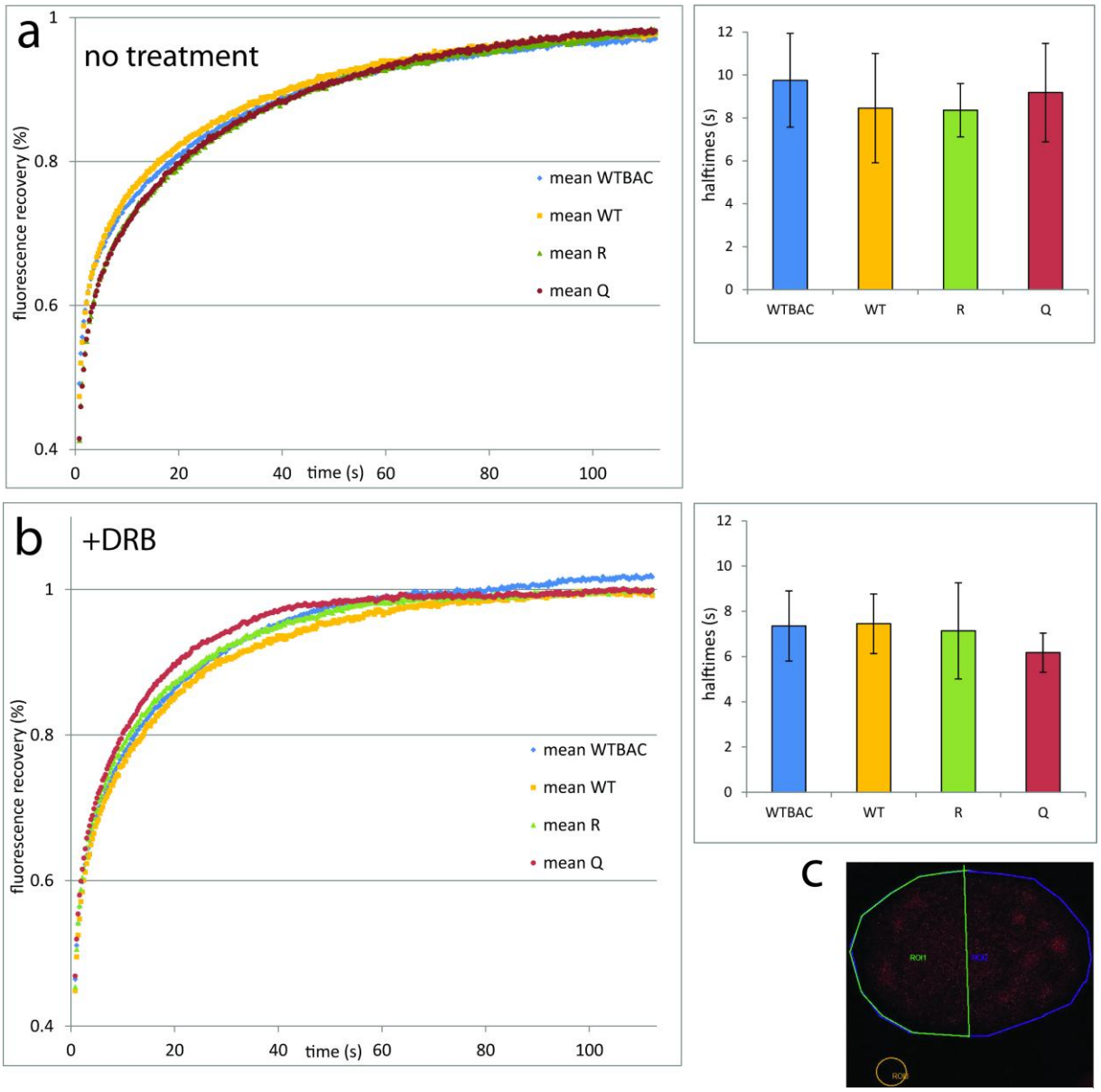

Fig. 6: (a) Average FRAP curves of SRSF5-GFP proteins and calculated half times of fluorescence recovery, both average of 8 to 12 cells, error bars represent standard deviation **(b)** The same measurements as previously but after inhibition of transcription by DRB **(c)** In the FRAP experiments half of the nucleus was bleached (green) the fluorescence was normalized to whole cell fluorescence (blue) and background was subtracted (orange).

To confirm that SRSF5-GFP proteins bind to RNA *in vivo*, I treated cells with DRB, an inhibitor of transcription elongation, for 4 hours thus depleting the nucleus of both pre-mRNA and mRNA and observed the protein dynamics again. In such condition pure diffusion of the proteins is observed. In all four cases the proteins moved faster and the half time of recovery shortened (**Fig. 6b**). In case of WTBAC and Q proteins the difference was statistically significant (p=0,007 and 0,002 respectively), unlike the WT and R, but the trend was the same in all four cases. The difference in half times with and without DRB was quite small, but the SRSF5 proteins were very fast even before the treatment, which means they interact shortly with other nuclear factors. FRAP results showed that SRSF5-GFP proteins

indeed bind to RNA and form transient interactions with it, but did not detect difference in RNA binding among them.

Quantification of SRSF5 interaction with RNA by RNA immunoprecipitation (RIP)

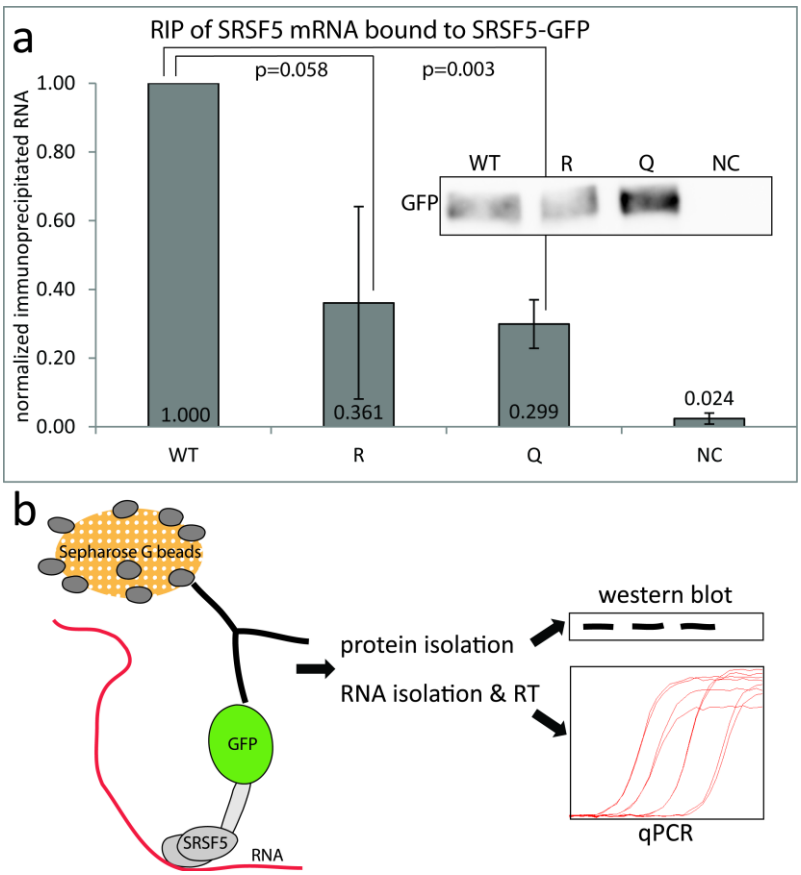

Fig. 7: (a) The amount of SRSF5 mRNA pulled down by SRSF5-GFP quantified as percent of input and normalized to amount of protein immunoprecipitated inferred from western blot. The WT value was set to 1. NC represents negative control,the HeLa cells not expressing GFP. Average of three biological experiments is shown, error bars represent standard deviation. Only a representative western blot from one experiment is shown **(b)** Scheme of the RIP experiment.

I decided to measure SRSF5 interaction with RNA in cell lysates by RIP. I followed a native protocol without crosslinking. It captures the situation *in vivo* quite faithfully, although the interactions can still change during the lysate incubation with antibody. I used the EGFP tag for immunoprecipitation and after the immunoprecipitation step, I isolated RNA from the immunoprecipitates and performed qRT-PCR (**Fig. 7b**). I quantified the RNA as percent of input. I also performed western blot with the protein fraction to quantify how much protein I pulled down, which was important for normalization of the results (**Fig. 7a**). From the WT and R cell lines I immunoprecipitated almost identical amount of SRSF5-EGFP, but from Q cells it was about 1.5 times more of the construct. This difference was not due to expression in the cell lines. It was rather the R cell line that had the highest expression of SRSF5-EGFP (**Fig. 4** and **9**), but for unknown reason the immunprecipitation was most efficient in the Q cell line. First, I optimized the protocol for better signal-to-noise ratio and then I performed three biological replicates. Few RNA targets of SRSF5 were known at the time and most of them were not sufficiently expressed

in HeLa cells to allow precise quantification by qPCR. So I measured SRSF5 binding to one target. It was an alternative exon in its own mRNA whose inclusion triggers nonsense-mediated decay (Lareau et al., 2007).

The mutant proteins R and Q pulled down approximately one third of SRSF5 mRNA as the WT. That was a first difference between WT and mutants observed. Neither the localization nor the dynamics of the proteins were able to detect such difference in RNA binding.

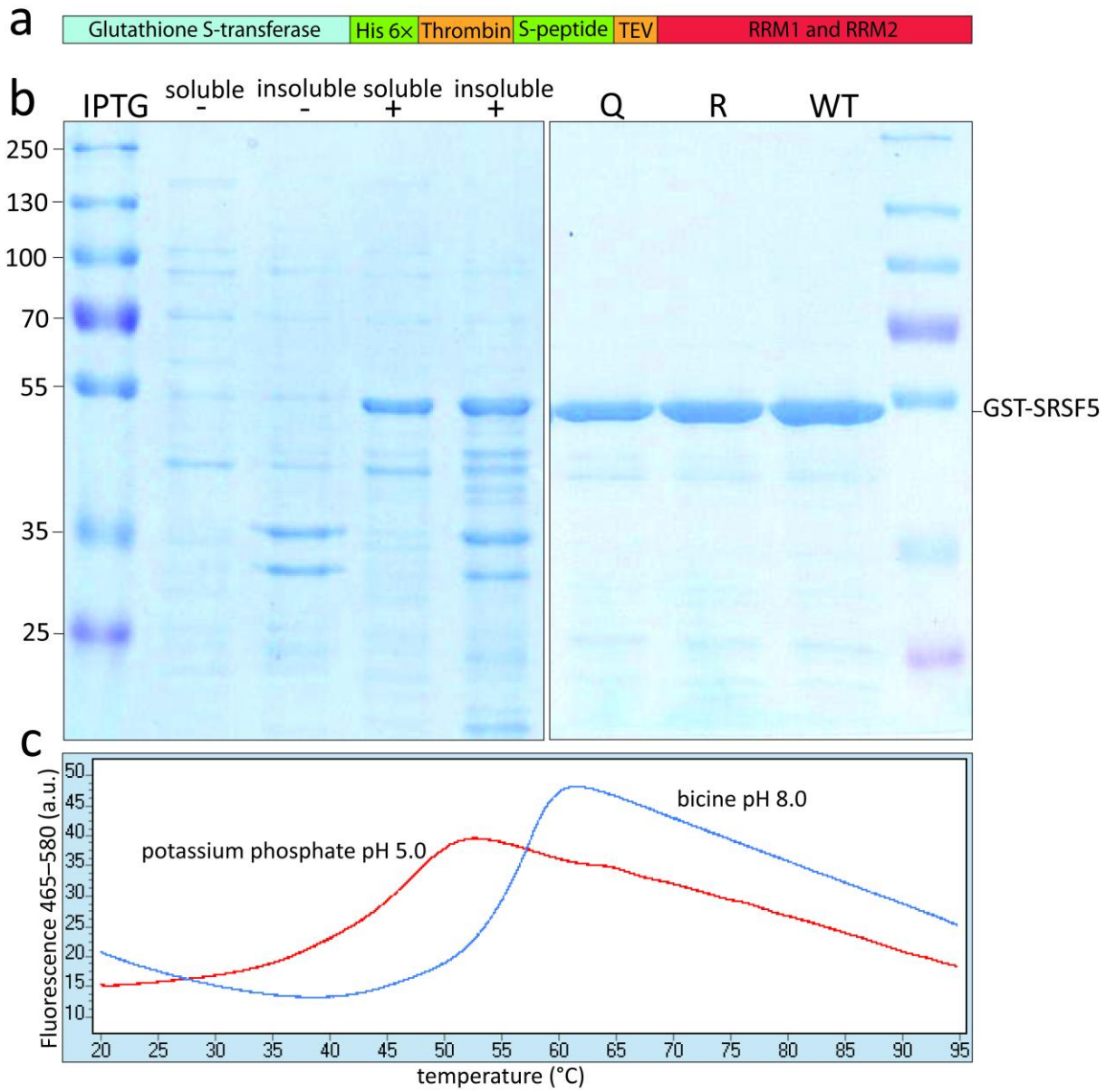

Fig. 8: (a) Scheme of the GST construct, His 6× is His-tag, Thrombin and TEV indicate protease cleavage sites and RRM1 and RRM2 denote the RNA binding domains of SRSF (first 180 amino acids). scheme is not to scale **(b)** On the left, SDS-PAGE of bacterial lysates before and after induction of GST-SRSF5 expression by IPTG, soluble and insoluble denote the supernatant and pelet after cell lysis by sonication, all three proteins behaved in the same way. on the right, purified GST-SRSF5 proteins. **(c)** An example of curves acquired by thermofluor measurement, bicine pH 8.0 was the best performing buffering agent.

GST-SRSF5 protein purification and buffer optimization

I decided to study binding of SRSF5 to RNA *in vitro* as well. For *in vitro* experiments I created another set of constructs containing the two RRM domains (amino acids 1–180) of WT, R and Q tagged on N-terminus with glutathione S-transferase (GST). The following construct had a size of 54.6 kDa and pI 8.3 (**Fig. 8a**). I expressed the fusion protein in bacteria. It was partially soluble (**Fig. 8b**) so I used the GST tag for protein purification on glutathione agarose beads. I was able to achieve good purity after GST purification, only some shorter, partially degraded fusion proteins copurified with the full length GST-SRSF5 (**Fig. 8b**).

	melting temperature °C			
buffers	R	R	Q	Q
NaCl	200	400	200	400
Sodium acetate 4.5	37,4	34,6	38,4	34,9
Sodium citrate 4.7	40,9	38,4	41,8	39,6
Sodium acetate 5.0	45,3	42,9	45,4	43,3
Potassium phosphate 5.0	47,1	35,5	47,9	46
Sodium phosphate 5.5	51,9	49,1	51,3	40
Sodium citrate 5.5	49,5	48,5	50,1	48,8
Mes 5.8	50,3	48,9	50,6	48,2
Potassium phosphate 6.0	51,3	50,1	51,6	50,2
Mes 6.2	52,6	49,6	52,6	50,7
Sodium phosphate 6.5	54,5	53	53,8	53,2
Mes 6.5	54	52,1	53,6	51,9
Potassium phosphate 7.0	54,2	53	53,8	52
Hepes 7.0	54,2	52,4	54,4	53
Sodium phosphate 7.5	54,8	53,7	54,8	53,1
Ammonium acetate 7.3	54,5	52,9	54,8	53,1
Tris 7.5	53,4	52,2	53,1	52
Imidazole 8.0	53,1	51,8	53,5	51,9
Hepes 8.0	54,9	53,2	54,8	52,7
Tris 8.0	53,7	52,5	53,5	52,5
Bicine 8.0	**57,5**	53,4	54,8	53
Tris 8.5	54,8	53,1	54,9	53
Bicine 9.0	55,7	53,8	**56,1**	53,5
ddH2O	53,3	53,2	53,1	52,6
Tris pH 8.0 300mM NaCl		54,8		54,8

Table 4: Thermofluor measurement of melting temperatures of R and Q proteins in 24 different buffers containing either 200 or 400 mM NaCl. Tris pH 8.0 300 mM NaCl is the original buffer proteins were eluted into.

After the purification I encountered problems with protein precipitation at higher concentration, so I performed a thermofluor measurement to find an optimal buffer for storage of purified GST-SRSF5 (**Fig. 8c**). I used the R and Q proteins because they differed by charge to see if each of the proteins needs unique conditions. I tested a panel of buffering agents all at 100 mM concentration with varying pH. I found that GST-SRSF5 is stable in alkaline conditions (**Table 4**) and the best buffer was bicine. Both R and Q proteins were similarly stable. Based on the thermofluor data I formulated the final buffer for storage and *in vitro* RNA binding assay as 100 mM bicine buffer (pH 7.6) containing 300 mM NaCl, 0.5 mM DTT and 10% glycerol. I decreased the pH from 8 to 7.6 because the in vitro assay was performed at low temperature and pH of the buffer prepared at room temperature increases at lower temperature.

Quantification of SRSF5 interaction with RNA *in vitro*

From the range of available assays for RNA-protein interaction, I opted for fluorescence anisotropy measurement. It was performed in the lab of Richard Štefl (CEITEC, Brno) by Veronika Bačíková. In this experiment the solution of 5′ fluorescein-labeled 12 bp long RNA 5′–UGACUCCACUGC–3′ consisting of two consecutive SRSF5 binding sites was titrated with purified GST-SRSF5(1–180) and increase in anisotropy was measured. The results were fitted, which allowed calculation of dissociation constant (K_d) of the RNA-protein interaction (**Fig. 9**).

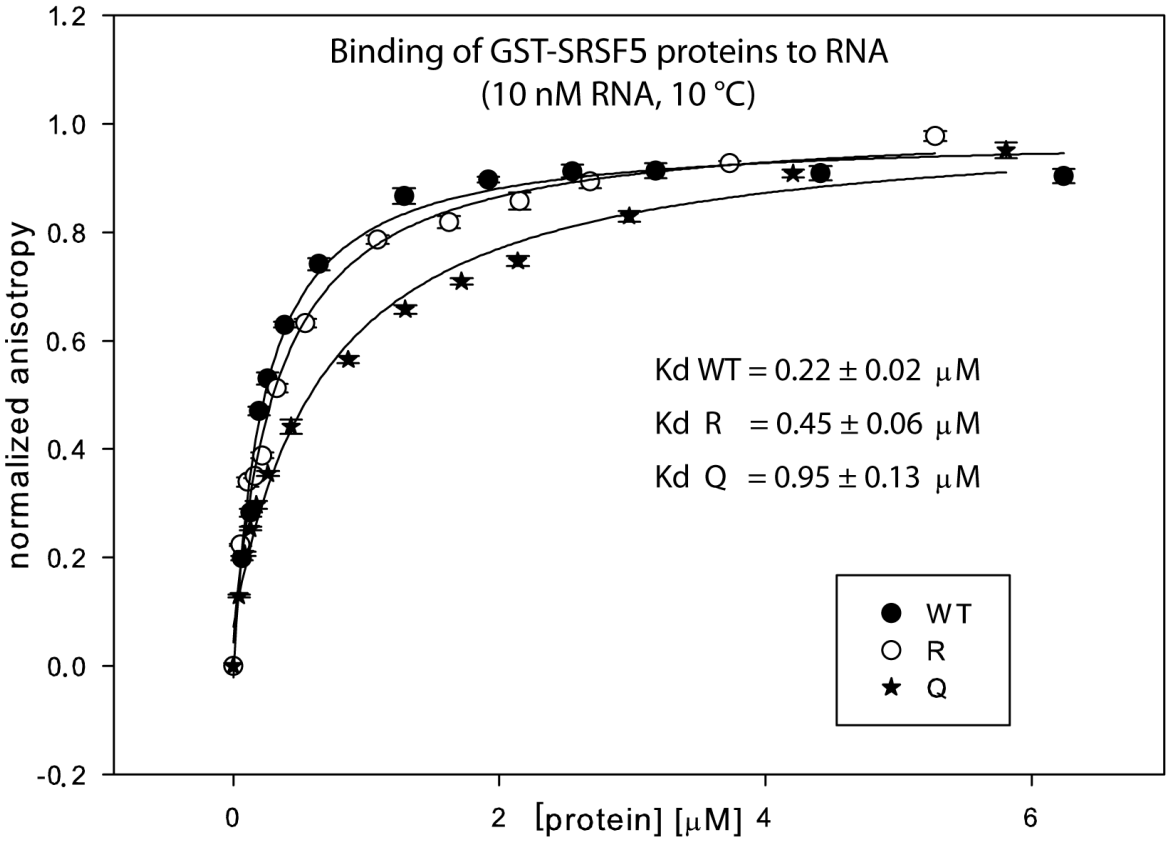

Fig. 9: Fluorescence anisotropy measurement and calculated dissociation constants (K_d) for WT, R and Q interaction with RNA.

WT protein had the lowest K_d of 0.22 μM followed by R 0.45 μM and Q protein with 0.95 μM. These results agree with the RIP measurements, where the mutant proteins bound less RNA than the WT. To control that the GST tag does not contribute to the binding of GST-SRSF5 to RNA, interaction of purified GST alone with the RNA probe was measured and it was confirmed that it does not bind RNA. To check the specificity of interaction of GST-SRSF5 with the designed probe containing SRSF5 binding sites, four other control RNA probes of identical length were measured with the WT protein. Their K_d varied and in two cases it was low, probably due to resemblance of the probe to the SRSF5 binding site and in two other

cases their interaction with WT protein was weak. Negative control no.1 5′–GGGGCCCCGUAA–3′ had the lowest affinity to SRSF5 with K_d of 5.27 μM. Negative control no.2 5′–GGGGCCCCCCCC–3′ had a K_d of 0.53 μM, negative control no.3 5′–GUAACCCCGGGG–3′ 0.51 μM and negative control no.4 5′–AAUCUUAGUAAUC–3′ bound RNA again weakly with K_d of 1.81 μM.

Influence of SRSF5 variants on splicing

I was ultimately interested in how WT and mutant SRSF5 proteins function in pre-mRNA splicing. To test SRSF5 variants in a functional assay, I used a splicing reporter based on human fibronectin gene that was used in our lab previously (Hnilicova et al., 2011). It contains a part of the gene from exon 24 to exon 26, where exon 25 (called EDB or EDII) is alternatively spliced and its inclusion is enhanced by SRSF5 binding directly to EDB exon and also to the intron downstream (White et al., 2008).

I attempted to knockdown the endogenous SRSF5 in the cell lines so that the SRSF5-EGFP is the only version of the protein in the cell. In such situation even a recessive effect on splicing is unobscured by presence of the functional endogenous protein. I used siRNA targeted against 3′ UTR of human SRSF5, but the knockdown was inefficient (**Fig. 10a,b**). I detected a decrease in mRNA level of endogenous SRSF5, but it was not paralleled by a decrease in protein level. Amount of mouse SRSF5-EGFP mRNA was variable, and surprisingly EGFP-tagged proteins accumulated afer knockdown of the endogenous SRSF5. I also checked that expression of WT, R or Q proteins does not change expression of other SR proteins (**Fig. 10a**). Even though the knockdown was not successfull I performed the splicing assay to see if the mutant proteins have a dominant effect on splicing (**Fig. 10c**). The splicing reporter produces two variants of mRNA. A short variant without EDB exon and a long one with EDB included. I observed the ratio of the short to long variant by RT-PCR. There was little difference in fibronectin reporter splicing among the tested cell lines (**Fig. 10c**).

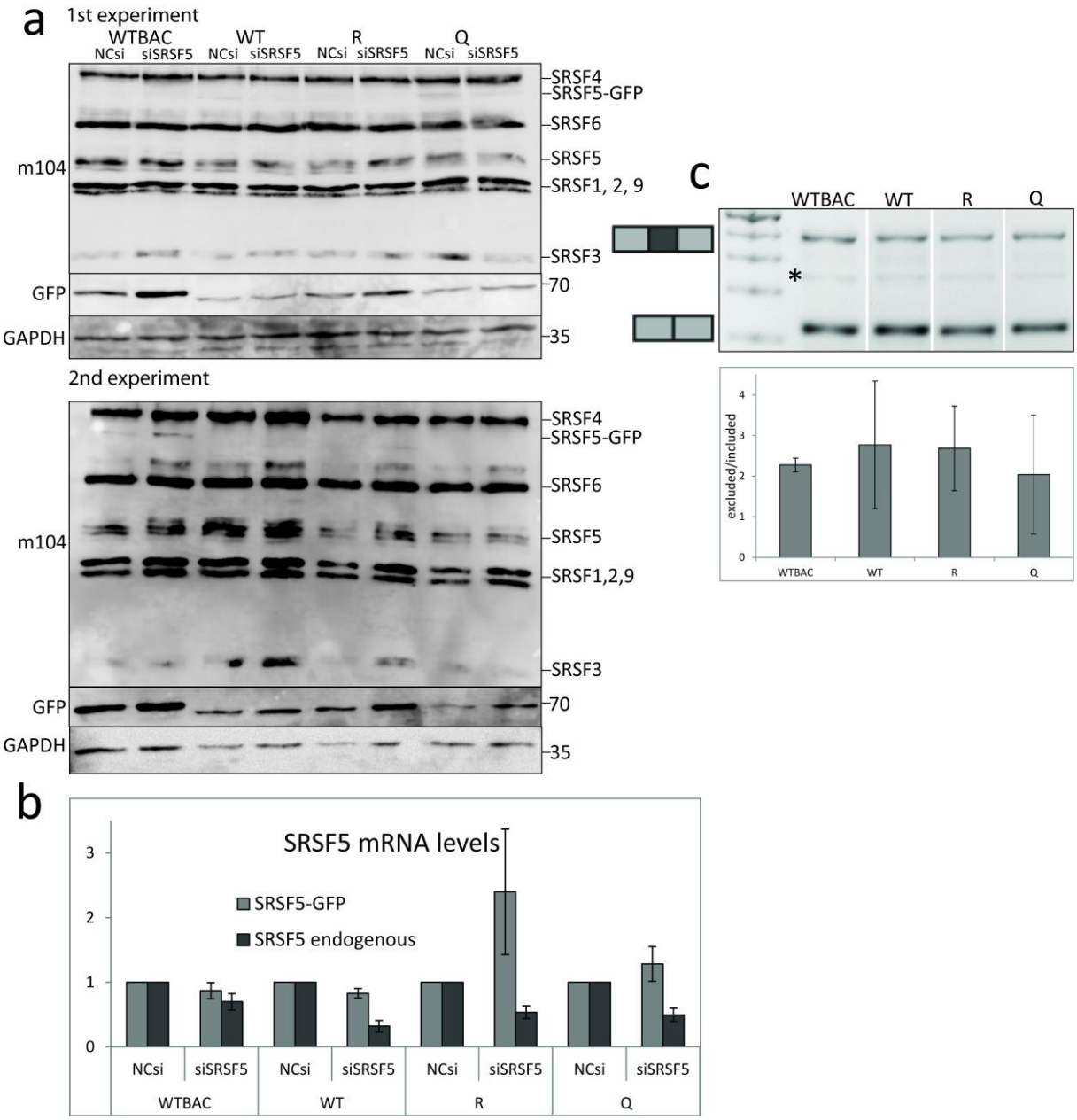

Fig. 10: (a) The changes in protein levels after knockdown of endogenous SRSF5. The level of endogenous SRSF5 probed with m104 antibody, SRSF5-GFP probed with GFP antibody, GAPDH served as a loading control. NCsi denotes use of negative control siRNA and siSRSF5 denotes siRNA against SRSF5. Two biological experiments are shown. **(b)** SRSF5 mRNA levels after knockdown measured by qPCR, average of the above-shown experiments is shown, error bars represent standard deviation. **(c)** DNA electrophoresis showing a ratio between short and long mRNA product from the fibronectin splicing reporter, below the denzitometric analysis in ImageJ, the band marked with asterisk is a hybrid between the short and long product and was neglected in the analysis, note that the image of the gel was cut and position of the lanes was changed for greater clarity.

Discussion

In this work I studied the influence of mutations in acetylation site on function of an SR protein SRSF5. I established cell lines expressing WT and mutant SRSF5 tagged with EGFP and thoroughly assessed their phenotype. First the localization of the constructs in cells was in good agreement with literature (Shepard and Hertel, 2009). They accumulated in nuclear speckles, which is typical for splicing factors and especially SR proteins. The proteins rearranged rapidly in response to transcription inhibition and also during the cell cycle proving that availability of their substrate RNA is one of determinants of their localization within nucleus.

Regarding the posttranslational modifications, I surprisingly detected the acetylation only in the wild type SRSF5 expressed from BAC (WTBAC) that served as a positive control and not in the wild type SRSF5 I cloned from cDNA. They differ mainly at RNA level because WTBAC contains its own promoter, UTRs and introns, whereas WT contains only the coding sequence. On the protein level they are very similar. Only WTBAC has longer linker sequence between SRSF5 and EGFP containing S-peptide and cleavage sites for TEV and Prescission proteases (Poser et al., 2008). It can be speculated that UTRs or introns are important for protein acetylation. But linking non-coding sequences in RNA to acetylation is problematic, because ε-N-acetylation of lysines occurs largely posttranslationally (unlike the N-terminal acetylation) and currently there is no known pathway that would transmit the information about presence of regulatory sequences in RNA onto the protein. Alternatively, acetylated lysine detected in WTBAC could be located in the linker sequence that is missing in the WT construct and does not reflect acetylation of SRSF5 itself.

The third possibility may be the crosstalk between acetylation and phosphorylation, because I revealed that the WTBAC is almost three times more phosphorylated than WT and mutant proteins. So WT might not be phosphorylated enough to get acetylated. Phosphorylation might be crucial for interaction with HAT or other proteins in the HAT complex and therefore essential for the acetylation. But from the western blot it cannot be deduced if the same subpopulation of SRSF5 undergoes both acetylation and phosphorylation, if the modifications are mutually exclusive or completely independent. For SRSF2 there exists a crosstalk between modifications as mentioned already in the introduction, but it is rather indirect. Overexpression of SRSF2 acetyltransferase Tip60 caused relocalization of SR protein kinases SRPK1 and SRPK2 from nucleus to cytoplasm and

thus decreased SRSF2 phosphorylation. The knockdown of Tip60 by siRNA had the opposite effect. On the other hand the SRSF2 deacetylase HDAC6 had no effect on SRSF2 kinases and SRSF5 phosphorylation (Edmond et al., 2011b).

Next experiment where I assessed the phenotype of SRSF5 mutants was fluorescence recovery after photobleaching. I was not able to detect differential interaction of SRSF5 proteins with RNA *in vivo* by this method, but I confirmed that all of them bind RNA because their recovery was faster after removal of binidng sites by transcription inhibition.

The fact, that I did not detect a difference in localization or dynamics between WTBAC, which is more phosphorylated and acetylated and WT and mutants, which are not, shows that other factors than posttranslational modifications are more significant for SRSF5 localization and dynamics. It was quite surprising in case of phosphorylation, because it is well established that it regulates SR protein localization (Misteli et al., 1998). The phosphorylated form should be nucleoplasmic and participate in splicing. So theoretically an increase in EGFP signal in nucleoplasm should have been detected in WTBAC, but a subtle change might have been neglected. The hypophosphorylated form of SR proteins tends to accumulate in speckles and also in cytoplasm, but such a pattern of EGFP signal was not observed in WT, R or Q cell lines.

The major hypothesis tested in this work was how efficiently do the SRSF5 proteins with mutated acetylation site (K167) bind RNA. I was working with wild type protein that had a positively charged lysine at position 167. This lysine was not acetylated in the cells (**Fig. 5**). I also assumed that proteins isolated from bacteria for *in vitro* assay did not carry any posttranslational modifications. Similar to the wild type was the R mutant with arginine instead of lysine at position 167. More dissimilar was the second Q mutant with lysine switched for glutamine, which lacked the positive charge and mimicked acetylated lysine in the experiments.

In the RNA immunoprecipitation experiment the three proteins were tested in their ability to pull down SRSF5 mRNA from the cell lysates. The mutants immunoprecipitated one third of the RNA pulled down by WT protein. The R mutant pulled down slightly more RNA than Q, but due to big variance of R values it is hard to conclude more. The fact that substitution K167R changed RNA binding properties similarly as the K167Q was unexpected because WT and R shared the positive charge at amino acid 167. This residue very probably plays a role in RNA-protein interaction, because any change at this position leads to

decreased RNA binding. This is further supported by analysis of Q mutant that interacted with RNA *in vitro* four times less efficiently than WT.

The *in vitro* RNA binding assay represented a simple system, where the number of variables was minimized. In these conditions the WT protein showed again the strongest binding to RNA with K_d= 0.22 µM followed by R with K_d= 0.45 µM and Q with K_d= 0.95 µM. This result confirmed the RIP data. From the K_d values it can be deduced that the positive charge at position 167 *in vitro* improves RNA binding. It agrees with the most straightforward model where positively charged amino acid residues interact with negatively charged phosphates in RNA. As mentioned in the introduction it is not always this way, but in case of SRSF5 acetylation probably decreases affinity to RNA.

None of the mutants failed completely to bind RNA which suggests that we should not think about acetylation as an inhibitor, but rather a modulator of RNA binding, which optimizes strength of the interaction, because the strongest ineraction of protein with RNA does not necessarily mean the best. Often weaker interaction is preferred, because not only binding, but also dissociation from the substrate is crucial for the reaction to proceed. Acetylation might be the decisive factor in alternative splicing events, which are susceptible to regulation and can easily come to different results.

The lower affinity of mutant proteins to RNA was not revealed *in vivo* by their different localisation or faster recovery in FRAP. The reason why the expected difference is not detected *in vivo* is probably caused by different conditions of RNA-protein interaction *in vitro* and *in vivo*, where it occurs in huge ribonucleoprotein particles (RNPs). RNA binding depends on a lot more macromolecular interactions and lower affinity to RNA is balanced by involvement of other protein-protein interactions of unreduced strength. So the major determinant of localisation or dynamics of SRSF5 in vivo is most likely the RS domain, which is responsible for most protein-protein interactions and outweighs the influence of RNA binding by RRM domain.

The fundamental question is not how acetylation influences binding of a protein to RNA, but how the altered RNA binding changes functioning of the protein. Therefore I decided to compare SRSF5 proteins in their influence on alternative splicing. Because the binding of SR proteins to splicing enhancer elements is at the core of their function, the assumption was that if the SR protein cannot bind to enhancer element, the exon in question should be more skipped. Such effect could not manifest itself on the background of functional endogenous

protein so I had to knock it down. Unfortunately the knockdown was not successfull and I saw only decrease in mRNA level, but the amount of endogenous human protein remained unchanged. In this experiment I observed an unexpected phenomenon. The decrease of mRNA of endogenous SRSF5 caused increase in expression of exogeous mouse protein. So there must be a mechanism that increases the stability or translation of exogenous SRSF5. SR proteins are known to have an intricate system of maintaining the level of their RNA present in cells by nonsense-mediated decay and clearly their regulation is even more complicated with additional mechanism at protein level documented here. This is not so surprising, because it is known that SR proteins also play a role in translation regulation.

The next step in the research of SRSF5 acetylation would be to abandon the mutants and study the acetylation of SRSF5 directly. One would need to know which HAT acetylates and which HDAC deacetylates SRSF5. It would enable two in principle similar experiments. SRSF5 could be acetylated *in vitro* and RNA binding assay *in vitro* could be repeated, which would alow direct comparison of RNA binding between acetylated and non-acetylated form of the protein. Second, by overexpression or knockdown of those enzymes one could regulate SRSF5 acetylation and observe how the properties of protein are changing *in vivo*. HAT and HDAC of SRSF2 were mapped, as mentioned previously. It is Tip60 and HDAC6 (Edmond et al., 2011b). Those would be first candidates for SRSF5 too, but it needs experimental validation.

This study could benefit from employment of mass spectrometry in several control experiments. There is a possibility that under normal conditions only small portion of protein is acetylated and phosphorylated, but this subpopulation gave the positive signal on western blot, however it is negligible when we observe the whole population of SRSF5 in cells. This would explain the lack of difference between mutants and WTBAC in localization and dynamics. This might be a concern for acetylation, which was hard to detect, but that is obviously influenced mostly by quality of antibody used. The ratio of acetylated to nonacetylated SRSF5 in cells can be determined only by mass spectrometry. This is not the only interesting question. We can ask also how this ratio changes in various conditions and by certain treatments. If high ratio of SRSF5 was acetylated, it would suggest that it has some functional significance. If little SRSF5 was acetylated or if the acetylation did not change in any condition, one would speculate that it has no functional significance and SRSF5 gets acetylated because some HAT evolved to acetylate similar epitope on a different

protein and acetylation of SRSF5 is only noise in the cellular acetylome. This might be a more general concern. The number of detected acetylations is huge and most of them still await confirmation in some functional study. Though the functional studies carried out so far confirmed that acetylation has a function in protein regulation, e.g., the metabolic enzymes (Zhao et al., 2010). However it can be caused by bias towards publishing positive results.

The high-throughput acetylation mass spectrometry data (Choudhary et al., 2009) also require validation. It has to be checked by mass spectrometry analysis of immunoprecipitated WTBAC protein that K167 is indeed acetylated. Additional acetylation sites might yet be discovered, because possibly some weak acetylation signal even in the R mutant was observed (**Fig. 5**).

Conclusions

I confirmed the hypothesis that the amino acid side chain at acetylation site (K167) in SRSF5 RRM2 is important for RNA binding. The mutant with lysine replaced by arginine bound RNA with affinity lowered to 36% *in vivo* or to one half *in vitro* compared with wild type according to two alternative RNA-protein interaction assays. So a conservative substitution preserving a positive charge at position 167 reduced RNA binding. The acetylation-mimicking mutant had even more decreased RNA affinity. To 30% *in vivo* or one quarter *in vitro* according to the two assays. This shows the importance of positive charge at position 167 for high-affinity RNA binding and suggests that acetylation could reduce SRSF5 binding to RNA.

Reduced binding of mutants to RNA did not manifest in their other properties like subcellular localization and dynamics and they were very probably functional. It remains to be tested if the reduced ability of mutant proteins to bind RNA influences pre-mRNA splicing and other functions of SRSF5. In the future, the direct effect of SRSF5 acetylation on RNA binding should be tested.

Literature

1. Allfrey, V.G., Faulkner, R., and Mirsky, A.E. (1964). Acetylation and Methylation of Histones and Their Possible Role in the Regulation of Rna Synthesis. Proc Natl Acad Sci U S A *51*, 786-794.
2. Anko, M.L., Morales, L., Henry, I., Beyer, A., and Neugebauer, K.M. (2010). Global analysis reveals SRp20-and SRp75-specific mRNPs in cycling and neural cells. Nature Structural & Molecular Biology *17*, 962-U973.
3. Anko, M.L., Muller-McNicoll, M., Brandl, H., Curk, T., Gorup, C., Henry, I., Ule, J., and Neugebauer, K.M. (2012). The RNA-binding landscapes of two SR proteins reveal unique functions and binding to diverse RNA classes. Genome Biol *13*, R17.
4. Anko, M.L., and Neugebauer, K.M. (2012). RNA-protein interactions in vivo: global gets specific. Trends Biochem Sci *37*, 255-262.
5. Arnesen, T. (2011). Towards a functional understanding of protein N-terminal acetylation. PLoS Biol *9*, e1001074.
6. Babic, I., Jakymiw, A., and Fujita, D.J. (2004). The RNA binding protein Sam68 is acetylated in tumor cell lines, and its acetylation correlates with enhanced RNA binding activity. Oncogene *23*, 3781-3789.
7. Bantscheff, M., Hopf, C., Savitski, M.M., Dittmann, A., Grandi, P., Michon, A.M., Schlegl, J., Abraham, Y., Becher, I., Bergamini, G.*, et al.* (2011). Chemoproteomics profiling of HDAC inhibitors reveals selective targeting of HDAC complexes. Nat Biotechnol *29*, 255-265.
8. Boyes, J., Byfield, P., Nakatani, Y., and Ogryzko, V. (1998). Regulation of activity of the transcription factor GATA-1 by acetylation. Nature *396*, 594-598.
9. Buscaino, A., Kocher, T., Kind, J.H., Holz, H., Taipale, M., Wagner, K., Wilm, M., and Akhtar, A. (2003). MOF-regulated acetylation of MSL-3 in the Drosophila dosage compensation complex. Mol Cell *11*, 1265-1277.
10. Busch, A., and Hertel, K.J. (2012). Evolution of SR protein and hnRNP splicing regulatory factors. Wiley Interdiscip Rev RNA *3*, 1-12.
11. Caputi, M., Freund, M., Kammler, S., Asang, C., and Schaal, H. (2004). A bidirectional SF2/ASF- and SRp40-dependent splicing enhancer regulates human immunodeficiency virus type 1 rev, env, vpu, and nef gene expression. J Virol *78*, 6517-6526.
12. Cartegni, L., Chew, S.L., and Krainer, A.R. (2002). Listening to silence and understanding nonsense: exonic mutations that affect splicing. Nat Rev Genet *3*, 285-298.
13. Cleary, J., Sitwala, K.V., Khodadoust, M.S., Kwok, R.P., Mor-Vaknin, N., Cebrat, M., Cole, P.A., and Markovitz, D.M. (2005). p300/CBP-associated factor drives DEK into interchromatin granule clusters. J Biol Chem *280*, 31760-31767.
14. Clery, A., Jayne, S., Benderska, N., Dominguez, C., Stamm, S., and Allain, F.H. (2011). Molecular basis of purine-rich RNA recognition by the human SR-like protein Tra2-beta1. Nat Struct Mol Biol *18*, 443-450.
15. D'Orso, I., and Frankel, A.D. (2009). Tat acetylation modulates assembly of a viral-host RNA-protein transcription complex. Proc Natl Acad Sci U S A *106*, 3101-3106.
16. Delcuve, G.P., Khan, D.H., and Davie, J.R. (2012). Roles of histone deacetylases in epigenetic regulation: emerging paradigms from studies with inhibitors. Clin Epigenetics *4*, 5.
17. Dominguez, C., and Allain, F.H. (2006). NMR structure of the three quasi RNA recognition motifs (qRRMs) of human hnRNP F and interaction studies with Bcl-x G-tract RNA: a novel mode of RNA recognition. Nucleic Acids Res *34*, 3634-3645.
18. Edmond, V., Brambilla, C., Brambilla, E., Gazzeri, S., and Eymin, B. (2011a). SRSF2 is required for sodium butyrate-mediated p21(WAF1) induction and premature senescence in human lung carcinoma cell lines. Cell Cycle *10*, 1968-1977.
19. Edmond, V., Moysan, E., Khochbin, S., Matthias, P., Brambilla, C., Brambilla, E., Gazzeri, S., and Eymin, B. (2011b). Acetylation and phosphorylation of SRSF2 control cell fate decision in response to cisplatin. EMBO J *30*, 510-523.

20. Ericsson, U.B., Hallberg, B.M., Detitta, G.T., Dekker, N., and Nordlund, P. (2006). Thermofluor-based high-throughput stability optimization of proteins for structural studies. Anal Biochem *357*, 289-298.
21. Garcia-Garcia, C., and Draper, D.E. (2003). Electrostatic interactions in a peptide--RNA complex. J Mol Biol *331*, 75-88.
22. Garrity, J., Gardner, J.G., Hawse, W., Wolberger, C., and Escalante-Semerena, J.C. (2007). N-lysine propionylation controls the activity of propionyl-CoA synthetase. J Biol Chem *282*, 30239-30245.
23. Goren, A., Ram, O., Amit, M., Keren, H., Lev-Maor, G., Vig, I., Pupko, T., and Ast, G. (2006). Comparative analysis identifies exonic splicing regulatory sequences--The complex definition of enhancers and silencers. Mol Cell *22*, 769-781.
24. Gu, W., and Roeder, R.G. (1997). Activation of p53 sequence-specific DNA binding by acetylation of the p53 C-terminal domain. Cell *90*, 595-606.
25. Guan, K.L., and Xiong, Y. (2011). Regulation of intermediary metabolism by protein acetylation. Trends Biochem Sci *36*, 108-116.
26. Guarente, L. (2011). The logic linking protein acetylation and metabolism. Cell Metab *14*, 151-153.
27. GuhaThakurta, D., and Draper, D.E. (2000). Contributions of basic residues to ribosomal protein L11 recognition of RNA. J Mol Biol *295*, 569-580.
28. Haberland, M., Montgomery, R.L., and Olson, E.N. (2009). The many roles of histone deacetylases in development and physiology: implications for disease and therapy. Nat Rev Genet *10*, 32-42.
29. Hargous, Y., Hautbergue, G.M., Tintaru, A.M., Skrisovska, L., Golovanov, A.P., Stevenin, J., Lian, L.Y., Wilson, S.A., and Allain, F.H. (2006). Molecular basis of RNA recognition and TAP binding by the SR proteins SRp20 and 9G8. EMBO J *25*, 5126-5137.
30. Hnilicova, J., Hozeifi, S., Duskova, E., Icha, J., Tomankova, T., and Stanek, D. (2011). Histone deacetylase activity modulates alternative splicing. PLoS One *6*, e16727.
31. Hottiger, M.O. (2011). ADP-ribosylation of histones by ARTD1: an additional module of the histone code? FEBS Lett *585*, 1595-1599.
32. Houtkooper, R.H., Pirinen, E., and Auwerx, J. (2012). Sirtuins as regulators of metabolism and healthspan. Nat Rev Mol Cell Biol *13*, 225-238.
33. Howitz, K.T., Bitterman, K.J., Cohen, H.Y., Lamming, D.W., Lavu, S., Wood, J.G., Zipkin, R.E., Chung, P., Kisielewski, A., Zhang, L.L*., et al.* (2003). Small molecule activators of sirtuins extend Saccharomyces cerevisiae lifespan. Nature *425*, 191-196.
34. Huangfu, D., Maehr, R., Guo, W., Eijkelenboom, A., Snitow, M., Chen, A.E., and Melton, D.A. (2008). Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. Nat Biotechnol *26*, 795-797.
35. Huo, L., Li, D., Sun, X., Shi, X., Karna, P., Yang, W., Liu, M., Qiao, W., Aneja, R., and Zhou, J. (2011). Regulation of Tat acetylation and transactivation activity by the microtubule-associated deacetylase HDAC6. J Biol Chem *286*, 9280-9286.
36. Huranova, M., Hnilicova, J., Fleischer, B., Cvackova, Z., and Stanek, D. (2009). A mutation linked to retinitis pigmentosa in HPRP31 causes protein instability and impairs its interactions with spliceosomal snRNPs. Hum Mol Genet *18*, 2014-2023.
37. Chen, J., and Li, Q. (2011). Life and death of transcriptional co-activator p300. Epigenetics *6*, 957-961.
38. Choudhary, C., Kumar, C., Gnad, F., Nielsen, M.L., Rehman, M., Walther, T.C., Olsen, J.V., and Mann, M. (2009). Lysine acetylation targets protein complexes and co-regulates major cellular functions. Science *325*, 834-840.
39. Imai, S., Armstrong, C.M., Kaeberlein, M., and Guarente, L. (2000). Transcriptional silencing and longevity protein Sir2 is an NAD-dependent histone deacetylase. Nature *403*, 795-800.
40. Jenuwein, T., and Allis, C.D. (2001). Translating the histone code. Science *293*, 1074-1080.

41. Jones, J.D., and O'Connor, C.D. (2011). Protein acetylation in prokaryotes. Proteomics *11*, 3012-3022.
42. Kaeberlein, M., McVey, M., and Guarente, L. (1999). The SIR2/3/4 complex and SIR2 alone promote longevity in Saccharomyces cerevisiae by two different mechanisms. Genes Dev *13*, 2570-2580.
43. Kalkhoven, E. (2004). CBP and p300: HATs for different occasions. Biochem Pharmacol *68*, 1145-1155.
44. Kiernan, R.E., Vanhulle, C., Schiltz, L., Adam, E., Xiao, H., Maudoux, F., Calomme, C., Burny, A., Nakatani, Y., Jeang, K.T*., et al.* (1999). HIV-1 tat transcriptional activity is regulated by acetylation. EMBO J *18*, 6106-6118.
45. Kim, J.W., and Dang, C.V. (2006). Cancer's molecular sweet tooth and the Warburg effect. Cancer Res *66*, 8927-8930.
46. Kim, S.C., Sprung, R., Chen, Y., Xu, Y., Ball, H., Pei, J., Cheng, T., Kho, Y., Xiao, H., Xiao, L*., et al.* (2006). Substrate and functional diversity of lysine acetylation revealed by a proteomics survey. Mol Cell *23*, 607-618.
47. Kouzarides, T. (2007). Chromatin modifications and their function. Cell *128*, 693-705.
48. Kuhn, A.N., van Santen, M.A., Schwienhorst, A., Urlaub, H., and Luhrmann, R. (2009). Stalling of spliceosome assembly at distinct stages by small-molecule inhibitors of protein acetylation and deacetylation. RNA *15*, 153-175.
49. L'Hernault, S.W., and Rosenbaum, J.L. (1985). Chlamydomonas alpha-tubulin is posttranslationally modified by acetylation on the epsilon-amino group of a lysine. Biochemistry *24*, 473-478.
50. Lareau, L.F., Inada, M., Green, R.E., Wengrod, J.C., and Brenner, S.E. (2007). Unproductive splicing of SR genes associated with highly conserved and ultraconserved DNA elements. Nature *446*, 926-929.
51. Law, M.J., Linde, M.E., Chambers, E.J., Oubridge, C., Katsamba, P.S., Nilsson, L., Haworth, I.S., and Laird-Offringa, I.A. (2006). The role of positively charged amino acids and electrostatic interactions in the complex of U1A protein and U1 hairpin II RNA. Nucleic Acids Res *34*, 275-285.
52. Liu, H.X., Zhang, M., and Krainer, A.R. (1998). Identification of functional exonic splicing enhancer motifs recognized by individual SR proteins. Genes Dev *12*, 1998-2012.
53. Long, J.C., and Caceres, J.F. (2009). The SR protein family of splicing factors: master regulators of gene expression. Biochem J *417*, 15-27.
54. Mahajan, R., Delphin, C., Guan, T., Gerace, L., and Melchior, F. (1997). A small ubiquitin-related polypeptide involved in targeting RanGAP1 to nuclear pore complex protein RanBP2. Cell *88*, 97-107.
55. Manley, J.L., and Krainer, A.R. (2010). A rational nomenclature for serine/arginine-rich protein splicing factors (SR proteins). Genes Dev *24*, 1073-1074.
56. Martinez-Balbas, M.A., Bauer, U.M., Nielsen, S.J., Brehm, A., and Kouzarides, T. (2000). Regulation of E2F1 activity by acetylation. EMBO J *19*, 662-671.
57. Misteli, T., Caceres, J.F., Clement, J.Q., Krainer, A.R., Wilkinson, M.F., and Spector, D.L. (1998). Serine phosphorylation of SR proteins is required for their recruitment to sites of transcription in vivo. J Cell Biol *143*, 297-307.
58. Nagamori, I., Cruickshank, V.A., and Sassone-Corsi, P. (2011). Regulation of an RNA granule during spermatogenesis: acetylation of MVH in the chromatoid body of germ cells. J Cell Sci *124*, 4346-4355.
59. Nagy, Z., and Tora, L. (2007). Distinct GCN5/PCAF-containing complexes function as co-activators and are involved in transcription factor and global histone acetylation. Oncogene *26*, 5341-5357.
60. Ni, J.Z., Grate, L., Donohue, J.P., Preston, C., Nobida, N., O'Brien, G., Shiue, L., Clark, T.A., Blume, J.E., and Ares, M., Jr. (2007). Ultraconserved elements are associated with

homeostatic control of splicing regulators by alternative splicing and nonsense-mediated decay. Genes Dev *21*, 708-718.
61. Norris, K.L., Lee, J.Y., and Yao, T.P. (2009). Acetylation goes global: the emergence of acetylation biology. Sci Signal *2*, pe76.
62. Pantoliano, M.W., Petrella, E.C., Kwasnoski, J.D., Lobanov, V.S., Myslik, J., Graf, E., Carver, T., Asel, E., Springer, B.A., Lane, P.*, et al.* (2001). High-density miniaturized thermal shift assays as a general strategy for drug discovery. J Biomol Screen *6*, 429-440.
63. Phair, R.D., Gorski, S.A., and Misteli, T. (2004a). Measurement of dynamic protein binding to chromatin in vivo, using photobleaching microscopy. Methods Enzymol *375*, 393-414.
64. Phair, R.D., Scaffidi, P., Elbi, C., Vecerova, J., Dey, A., Ozato, K., Brown, D.T., Hager, G., Bustin, M., and Misteli, T. (2004b). Global nature of dynamic protein-chromatin interactions in vivo: three-dimensional genome scanning and dynamic interaction networks of chromatin proteins. Mol Cell Biol *24*, 6393-6402.
65. Phelan, M.M., Goult, B.T., Clayton, J.C., Hautbergue, G.M., Wilson, S.A., and Lian, L.Y. (2012). The structure and selectivity of the SR protein SRSF2 RRM domain with RNA. Nucleic Acids Res *40*, 3232-3244.
66. Polesskaya, A., Duquet, A., Naguibneva, I., Weise, C., Vervisch, A., Bengal, E., Hucho, F., Robin, P., and Harel-Bellan, A. (2000). CREB-binding protein/p300 activates MyoD by acetylation. J Biol Chem *275*, 34359-34364.
67. Poser, I., Sarov, M., Hutchins, J.R., Heriche, J.K., Toyoda, Y., Pozniakovsky, A., Weigl, D., Nitzsche, A., Hegemann, B., Bird, A.W.*, et al.* (2008). BAC TransgeneOmics: a high-throughput method for exploration of protein function in mammals. Nat Methods *5*, 409-415.
68. Reddy, A.S., and Shad Ali, G. (2011). Plant serine/arginine-rich proteins: roles in precursor messenger RNA splicing, plant development, and stress responses. Wiley Interdiscip Rev RNA *2*, 875-889.
69. Sanford, J.R., Wang, X., Mort, M., Vanduyn, N., Cooper, D.N., Mooney, S.D., Edenberg, H.J., and Liu, Y. (2009). Splicing factor SFRS1 recognizes a functionally diverse landscape of RNA transcripts. Genome Res *19*, 381-394.
70. Sapountzi, V., and Cote, J. (2011). MYST-family histone acetyltransferases: beyond chromatin. Cell Mol Life Sci *68*, 1147-1156.
71. Scroggins, B.T., Robzyk, K., Wang, D., Marcu, M.G., Tsutsumi, S., Beebe, K., Cotter, R.J., Felts, S., Toft, D., Karnitz, L.*, et al.* (2007). An acetylation site in the middle domain of Hsp90 regulates chaperone function. Mol Cell *25*, 151-159.
72. Shepard, P.J., and Hertel, K.J. (2009). The SR protein family. Genome Biol *10*, 242.
73. Schwer, B., Eckersdorff, M., Li, Y., Silva, J.C., Fermin, D., Kurtev, M.V., Giallourakis, C., Comb, M.J., Alt, F.W., and Lombard, D.B. (2009). Calorie restriction alters mitochondrial protein acetylation. Aging Cell *8*, 604-606.
74. Spange, S., Wagner, T., Heinzel, T., and Kramer, O.H. (2009). Acetylation of non-histone proteins modulates cellular signalling at multiple levels. Int J Biochem Cell Biol *41*, 185-198.
75. Starai, V.J., Celic, I., Cole, R.N., Boeke, J.D., and Escalante-Semerena, J.C. (2002). Sir2-dependent activation of acetyl-CoA synthetase by deacetylation of active lysine. Science *298*, 2390-2392.
76. Sun, Y., Jiang, X., Chen, S., Fernandes, N., and Price, B.D. (2005). A role for the Tip60 histone acetyltransferase in the acetylation and activation of ATM. Proc Natl Acad Sci U S A *102*, 13182-13187.
77. Sykes, S.M., Mellert, H.S., Holbert, M.A., Li, K., Marmorstein, R., Lane, W.S., and McMahon, S.B. (2006). Acetylation of the p53 DNA-binding domain regulates apoptosis induction. Mol Cell *24*, 841-851.
78. Tacke, R., Chen, Y., and Manley, J.L. (1997). Sequence-specific RNA binding by an SR protein requires RS domain phosphorylation: creation of an SRp40-specific splicing enhancer. Proc Natl Acad Sci U S A *94*, 1148-1153.

79. Tang, X., Gao, J.S., Guan, Y.J., McLane, K.E., Yuan, Z.L., Ramratnam, B., and Chin, Y.E. (2007). Acetylation-dependent signal transduction for type I interferon receptor. Cell *131*, 93-105.
80. Tang, Y., Luo, J., Zhang, W., and Gu, W. (2006). Tip60-dependent acetylation of p53 modulates the decision between cell-cycle arrest and apoptosis. Mol Cell *24*, 827-839.
81. Thompson, P.R., Wang, D., Wang, L., Fulco, M., Pediconi, N., Zhang, D., An, W., Ge, Q., Roeder, R.G., Wong, J.*, et al.* (2004). Regulation of the p300 HAT domain via a novel activation loop. Nat Struct Mol Biol *11*, 308-315.
82. Tintaru, A.M., Hautbergue, G.M., Hounslow, A.M., Hung, M.L., Lian, L.Y., Craven, C.J., and Wilson, S.A. (2007). Structural and functional analysis of RNA and TAP binding to SF2/ASF. EMBO Rep *8*, 756-762.
83. Tripathi, V., Ellis, J.D., Shen, Z., Song, D.Y., Pan, Q., Watt, A.T., Freier, S.M., Bennett, C.F., Sharma, A., Bubulya, P.A.*, et al.* (2010). The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. Mol Cell *39*, 925-938.
84. Twyffels, L., Gueydan, C., and Kruys, V. (2011). Shuttling SR proteins: more than splicing factors. FEBS J *278*, 3246-3255.
85. Wang, C., Tian, L., Popov, V.M., and Pestell, R.G. (2011). Acetylation and nuclear receptor action. J Steroid Biochem Mol Biol *123*, 91-100.
86. Wang, L., Du, Y., Lu, M., and Li, T. (2012). ASEB: a web server for KAT-specific acetylation site prediction. Nucleic Acids Res.
87. Wang, Q., Zhang, Y., Yang, C., Xiong, H., Lin, Y., Yao, J., Li, H., Xie, L., Zhao, W., Yao, Y.*, et al.* (2010). Acetylation of metabolic enzymes coordinates carbon source utilization and metabolic flux. Science *327*, 1004-1007.
88. Wellen, K.E., and Thompson, C.B. (2012). A two-way street: reciprocal regulation of metabolism and signalling. Nat Rev Mol Cell Biol *13*, 270-276.
89. White, E.S., Baralle, F.E., and Muro, A.F. (2008). New insights into form and function of fibronectin splice variants. J Pathol *216*, 1-14.
90. Yang, X.J., and Seto, E. (2008). Lysine acetylation: codified crosstalk with other posttranslational modifications. Mol Cell *31*, 449-461.
91. Yu, B.J., Kim, J.A., Moon, J.H., Ryu, S.E., and Pan, J.G. (2008). The diversity of lysine-acetylated proteins in Escherichia coli. J Microbiol Biotechnol *18*, 1529-1536.
92. Zhang, J., Sprung, R., Pei, J., Tan, X., Kim, S., Zhu, H., Liu, C.F., Grishin, N.V., and Zhao, Y. (2009). Lysine acetylation is a highly abundant and evolutionarily conserved modification in Escherichia coli. Mol Cell Proteomics *8*, 215-225.
93. Zhang, Z., Tan, M., Xie, Z., Dai, L., Chen, Y., and Zhao, Y. (2011). Identification of lysine succinylation as a new post-translational modification. Nat Chem Biol *7*, 58-63.
94. Zhao, S., Xu, W., Jiang, W., Yu, W., Lin, Y., Zhang, T., Yao, J., Zhou, L., Zeng, Y., Li, H.*, et al.* (2010). Regulation of cellular metabolism by protein lysine acetylation. Science *327*, 1000-1004.
95. Zhong, X.Y., Wang, P., Han, J., Rosenfeld, M.G., and Fu, X.D. (2009). SR proteins in vertical integration of gene expression from transcription to RNA processing to translation. Mol Cell *35*, 1-10.
96. Zhou, Y., Schmitz, K.M., Mayer, C., Yuan, X., Akhtar, A., and Grummt, I. (2009). Reversible acetylation of the chromatin remodelling complex NoRC is required for non-coding RNA-dependent silencing. Nat Cell Biol *11*, 1010-1016.